I. Guikhman, A. Skorokhod

INTRODUCTION À LA THÉORIE DES PROCESSUS ALÉATOIRES

Éditions Mir Moscou

I. GUIKHMAN, A. SKOROKHOD

INTRODUCTION À LA THÉORIE DES PROCESSUS ALÉATOIRES

ÉDITIONS MIR · MOSCOU

Traduit du russe
par DJILALI EMBAREK

Le vocabulaire technique français du texte est raisonnable et compréhensible, mais n'est pas toujours celui utilisé en français. Ainsi, en suivant l'index terminologique, on a par exemple :

Absence de post-action [bien vu, mais non usité]
Condition de brassage pour condition de mélange
Flot de tribus pour filtration
Instant markovien pour temps d'arrêt
Instant de première atteinte pour temps d'entrée
Processus branchu pour processus de branchement
Submartingale pour sousmartingale
Supermartingale pour surmartingale
Tabou probabilité pour probabilité tabou

На французском языке

TABLE DES MATIÈRES

LISTE DES NOTATIONS

$\forall x \in X$: pour tous les $x \in X$,

$\exists x \in X$: il existe $x \in X$,

$\varnothing$: ensemble vide,

$A \subset B$: B contient A (l'événement A implique l'événement B),

$\bigcup$: union d'ensembles (somme d'événements),

$\bigcap$: intersection d'ensembles (superposition d'événements),

$A \setminus B$: différence des ensembles A et B,

$\overline{A}$: complémentaire de l'ensemble A,

$\{x : A\}$: ensemble des éléments x vérifiant la relation A,

χ_A ou $\chi(A)$: indicateur de l'événement (de l'ensemble) A,

$\mathscr{R}^d$: espace euclidien de dimension d,

$h \uparrow a$ ($h \downarrow a$) : h tend en croissant (resp. en décroissant) vers a,

(x, y) : produit scalaire des vecteurs x et y,

δ_{ij} : symbole de Kronecker égal à 1 pour $i = j$, et à 0 pour $i \neq j$,

$\vee\ b$ ($a \wedge b$) : le plus grand (le plus petit) des nombres a et b,

$\mathsf{P}\{A\}$: probabilité de l'événement A,

$\mathsf{E}\xi$: espérance mathématique de la variable ξ,

$\operatorname{Var} \xi$: variance de la variable ξ.

AVANT-PROPOS

Ce livre s'adresse aux personnes possédant les connaissances indispensables de la théorie des probabilités et désireuses d'entreprendre l'étude de la théorie des processus aléatoires. Les auteurs nourrissent l'espoir qu'il sera profitable aux élèves du second et du troisième cycle, ainsi qu'à tous ceux, non mathématiciens, qui voudraient s'initier aux principaux résultats et méthodes de la théorie dans une forme rigoureuse mais ni exhaustive, ni la plus générale.

Les auteurs ne se sont pas fixé pour objectif de traiter tous les chapitres de la théorie, ils ont omis certains problèmes et méthodes développés dans la littérature consacrée à ce sujet (semi-groupes dans la théorie des processus markoviens, propriétés ergodiques des processus markoviens, processus aléatoires généraux).

La théorie des processus aléatoires s'est détachée depuis peu de la théorie des probabilités. Cependant elle lui est si étroitement liée qu'il est souvent très malaisé de faire la démarcation. Ainsi la théorie des processus aléatoires se rattache à la théorie de sommation des variables aléatoires indépendantes par le chapitre consacré aux processus à accroissements indépendants, et à la statistique mathématique par les problèmes de statistique de la théorie des processus aléatoires.

Voici (à notre sens) les problèmes fondamentaux de la théorie des processus aléatoires.

1. La construction d'un modèle mathématique autorisant une définition rigoureuse (formelle) du processus aléatoire et l'étude de ses propriétés générales.

2. La classification des processus aléatoires. Toute classification étant un choix délibéré, il importera de se fonder sur des principes qui en indiqueraient au moins l'« orientation ». La classification adoptée met en évidence certaines classes admettant une définition plus ou moins constructive. Chaque classe est caractérisée par le fait qu'il suffit de se donner subsidiairement seulement un nombre fini de caractéristiques fonctionnelles pour en sélectionner un processus aléatoire particulier.

On étudie parfois des classes de processus admettant une même méthode de résolution. Dans ces cas, généralement, on ne s'intéresse pas aux aspects distinctifs des processus, si seulement coïncident les caractéristiques indispensables à la résolution de ces problèmes.

Signalons les classes suivantes de processus:

a) processus à accroissements indépendants,

b) processus markoviens,

c) processus gaussiens,

d) processus stationnaires au sens strict,

e) processus stationnaires au sens large (on peut leur ajouter les processus à accroissements stationnaires).

3. Le problème, étroitement rattaché au précédent, qui consiste à élaborer pour les diverses classes de processus aléatoires un outil mathématique permettant de calculer les caractéristiques probabilistes des processus aléatoires. Pour les caractéristiques élémentaires cet outil existe déjà et utilise en principe soit la théorie des équations différentielles (ordinaires et à dérivées partielles), soit la théorie des équations intégro-différentielles (pour les processus markoviens et les processus à accroissements indépendants), soit la théorie des équations intégrales à noyaux symétriques (si les processus sont gaussiens), soit enfin la transformation de Fourier et la théorie des fonctions d'une variable complexe (pour les processus à accroissements indépendants et les processus stationnaires).

4. Le problème, très important par sa valeur pratique et son rôle dans l'élaboration de certains chapitres de la théorie des processus aléatoires, qui, dans sa forme générale, consiste à déterminer la meilleure valeur d'une fonctionnelle de processus d'après les valeurs d'autres fonctionnelles du même processus. Exemple: le problème de prédiction qui consiste à déterminer la valeur d'un processus à une date quelconque non comprise dans l'intervalle sur lequel il est observé.

5. L'étude des diverses transformations des processus aléatoires. Ces transformations permettent de ramener les processus complexes à de plus simples. La théorie des équations différentielles et intégrales des processus aléatoires de même d'ailleurs que les théorèmes limites (qui s'apparentent à une transformation) sont traités dans le cadre des transformations de processus aléatoires.

Les méthodes de la théorie des processus aléatoires ont gagné pratiquement tous les domaines de la science.

La théorie des processus aléatoires intervient en radio et en électronique (processus stationnaires au sens large et processus gaussiens), en cybernétique (processus stationnaires au sens strict et processus markoviens) ainsi qu'en économie, biologie, physique, etc.

Voici les grands traits de cet ouvrage.

Le premier chapitre est consacré aux processus aléatoires au sens large, c'est-à-dire à cette partie de la théorie qui traite des répartitions d'ensembles finis de valeurs d'un processus aléatoire. Cette partie est proche de la théorie élémentaire des probabilités, elle n'implique pas de notions mathématiques complexes et suffit souvent aux applications.

On étudie ensuite l'axiomatique de la théorie des probabilités, certains problèmes de la théorie des processus aléatoires à temps accéléré (martingales, théorème ergodique, chaînes de Markov), les problèmes généraux de la théorie des fonctions aléatoires et ensuite des classes concrètes de processus aléatoires et des problèmes particuliers. Les processus aléatoires à accroissements indépendants (un chapitre) et les processus markoviens (deux chapitres) sont largement développés. Les processus stationnaires sont traités dans une partie du premier chapitre et dans une partie du cinquième, qui est consacré aux transformations linéaires des processus aléatoires. Ce chapitre développe également le problème de prédiction linéaire. Les théorèmes limites pour processus aléatoires, principalement les processus à accroissements indépendants et les processus markoviens, occupent tout un chapitre.

La plupart des constructions sont effectuées pour le cas où le processus aléatoire prend ses valeurs dans un espace euclidien de dimension finie. On envisage parfois des processus à une ou plusieurs dimensions à valeurs complexes dans un espace métrique complet.

On suppose au lecteur la connaissance des notions élémentaires d'algèbre linéaire (en particulier, pour l'étude des processus gaussiens), de théorie des espaces hilbertiens (pour les transformations linéaires de processus aléatoires) et d'analyse fonctionnelle (espace métrique complet, compacts).

Notre objectif n'a pas été de donner une bibliographie complète des ouvrages traitant de la théorie des processus aléatoires: outre les livres auxquels il est fait référence dans le texte, nous n'avons cité que les ouvrages fondamentaux de théorie des processus aléatoires et de théorie des probabilités existant en russe, ainsi que les articles exposant les premiers résultats fondamentaux dans le domaine considéré.

Le livre se divise en chapitres, les chapitres en paragraphes. Les formules importantes, les théorèmes et les lemmes sont numérotés dans le cadre de chaque paragraphe.

Les auteurs expriment leur gratitude aux chercheurs, aspirants et étudiants de la chaire de théorie des probabilités et de statistique mathématique de l'Université de Kiev pour le concours apporté à la conception de cet ouvrage.

I. Guikhman,
A. Skorokhod

CHAPITRE PREMIER

PROCESSUS ALÉATOIRES AU SENS LARGE

§ 1. Définitions

Un processus aléatoire, comme d'ailleurs un processus déterministe, se décrit par une fonction $\xi(\theta)$ (à valeurs réelles, complexes ou vectorielles), où θ, l'argument de la fonction, parcourt un ensemble Θ. La fonction $\xi(\theta)$ observée au cours d'une expérience assujettie à certaines conditions Y s'appelle *fonction échantillonnée* ou *réalisation* du processus aléatoire.

Si θ est fixé, la valeur $\xi(\theta)$ est aléatoire. Pour pouvoir appliquer les méthodes mathématiques aux problèmes étudiés, il est naturel de supposer que $\xi(\theta)$ est une variable aléatoire (ou un vecteur aléatoire) au sens probabiliste.

Un processus aléatoire est donc une famille de variables aléatoires $\xi(\theta)$ dépendant d'un paramètre $\theta \in \Theta$.

Si l'ensemble Θ est arbitraire, il est plus commode de remplacer le terme *processus aléatoire* par fonction aléatoire, réservant la dénomination *processus aléatoire* aux cas où le paramètre θ figure le temps. Si l'argument de la fonction aléatoire est une variable tridimensionnelle, cette fonction est dite *champ aléatoire*.

Cette définition de la fonction aléatoire appelle quelques précisions. Pour simplifier on n'envisagera que des fonctions aléatoires à valeurs réelles. Mais tout d'abord il importe d'expliciter ce qu'on entend par famille de variables aléatoires dépendant d'un paramètre θ. On rappelle qu'en théorie des probabilités, une suite finie de variables aléatoires $\xi_1, \xi_2, \ldots, \xi_n$ est entièrement caractérisée par la fonction de répartition conjointe de ces variables

$$F(x_1, x_2, \ldots, x_n) = \mathsf{P}\{\xi_1 < x_1, \xi_2 < x_2, \ldots, \xi_n < x_n\}.$$

La description probabiliste de la fonction aléatoire soulève le problème suivant: comment décrire les relations qui lient entre elles l'infinité des valeurs aléatoires prises par la fonction aléatoire?

Le plus simple est de considérer la fonction aléatoire $\xi(\theta)$ donnée si sont définies toutes les relations probabilistes liant une collection finie quelconque de valeurs prises par les variables aléatoires

$$\xi(\theta_1), \xi(\theta_2), \ldots, \xi(\theta_n), \theta_i \in \Theta, \quad (1)$$
$$i = 1, 2, \ldots, n; \; n = 1, 2, \ldots,$$

c'est-à-dire si sont données les fonctions de répartition correspondantes. Dans cette optique la fonction aléatoire $\xi(\theta)$, $\theta \in \Theta$, est définie par la famille de répartitions

$$F_{\theta_1, \theta_2, \ldots, \theta_n}(x_1, x_2, \ldots, x_n); \ \theta_i \in \Theta, \quad i = 1, 2, \ldots, n; \quad n = 1, 2, \ldots, \tag{2}$$

et chaque fonction $F_{\theta_1, \theta_2, \ldots, \theta_n}(x_1, x_2, \ldots, x_n)$ est interprétée comme la fonction de répartition conjointe de la suite de variables aléatoires (1).

Il va de soi que la validité de cette interprétation implique que la famille de répartitions (2) ne soit pas entièrement arbitraire. Elle doit en effet satisfaire les conditions évidentes suivantes appelées *conditions de compatibilité* de la famille de répartitions (2):

$$F_{\theta_1, \theta_2, \ldots, \theta_{n+1}, \ldots, \theta_{n+p}}(x_1, x_2, \ldots, x_n, +\infty, \ldots, +\infty) = F_{\theta_1, \theta_2, \ldots, \theta_n}(x_1, x_2, \ldots, x_n), \tag{3}$$

$$F_{\theta_1, \theta_2, \ldots, \theta_n}(x_1, x_2, \ldots, x_n) = F_{\theta_{i_1}, \theta_{i_2}, \ldots, \theta_{i_n}}(x_{i_1}, x_{i_2}, \ldots, x_{i_n}), \tag{4}$$

où $i_1, i_2, \ldots, i_n$ est une permutation quelconque des indices $1, 2, \ldots, n$.

Ces conditions sont nécessaires. En effet,

$$F_{\theta_1, \theta_2, \ldots, \theta_n, \theta_{n+1}, \ldots, \theta_{n+p}}(x_1, x_2, \ldots, x_n, +\infty, \ldots, +\infty) = \mathsf{P}\{\xi(\theta_1) < x_1, \xi(\theta_2) < x_2, \ldots, \xi(\theta_n) < x_n, \xi(\theta_{n+1}) < \infty, \ldots, \xi(\theta_{n+p}) < \infty\} = \mathsf{P}\{\xi(\theta_1) < x_1, \xi(\theta_2) < x_2, \ldots, \xi(\theta_n) < x_n\} = F_{\theta_1, \theta_2, \ldots, \theta_n}(x_1, x_2, \ldots, x_n),$$

$$F_{\theta_1, \theta_2, \ldots, \theta_n}(x_1, x_2, \ldots, x_n) = \mathsf{P}\{\xi(\theta_1) < x_1, \xi(\theta_2) < x_2, \ldots, \xi(\theta_n) < x_n\} = \mathsf{P}\{\xi(\theta_{i_1}) < x_{i_1}, \xi(\theta_{i_2}) < x_{i_2}, \ldots, \xi(\theta_{i_n}) < x_{i_n}\} = F_{\theta_{i_1}, \theta_{i_2}, \ldots, \theta_{i_n}}(x_{i_1}, x_{i_2}, \ldots, x_{i_n}).$$

Ce qui précède nous conduit à la définition suivante.

DÉFINITION. *On appelle fonction aléatoire* $\xi(\theta)$ *définie sur un ensemble* Θ ($\theta \in \Theta$) *et à valeurs réelles, une famille de répartitions* (2) *vérifiant les conditions de compatibilité* (3), (4).

Les fonctions $F_{\theta_1, \theta_2, \ldots, \theta_n}(x_1, x_2, \ldots, x_n)$ sont appelées *répartitions finidimensionnelles* de la fonction aléatoire.

Le point fort de la définition de la fonction aléatoire donnée plus haut est qu'elle est élémentaire et suffit dans les cas où l'on s'intéresse aux valeurs prises par les variables aléatoires $\xi(\theta)$ sur un ensemble fini de valeurs de l'argument θ. Son point faible est

qu'elle ne permet pas de considérer une fonction aléatoire dans son ensemble, c'est-à-dire de considérer simultanément l'ensemble de toutes ses valeurs. Cependant dans de nombreuses épreuves la fonction échantillonnée est décrite à l'aide d'un instrument approprié sous forme de graphe. La définition donnée non seulement ne permet pas de tracer ce graphe, mais elle ne permet pas non plus d'étudier les propriétés de la fonction $\xi(\theta)$ telles la continuité, la dérivabilité, etc., ni même la probabilité que l'événement $\{a < \xi(\theta) < b,\ a < b\}$ est réalisé $\forall\theta \in \Theta$.

L'approche axiomatique de la théorie des probabilités donne lieu à d'autres définitions plus élégantes de la fonction aléatoire. Tout schéma probabiliste décrit les résultats d'une expérience à issue aléatoire. Si le résultat est décrit par un seul nombre ou par une suite finie de nombres, on dit qu'on observe une variable aléatoire ou un vecteur aléatoire. Si par contre ce résultat est décrit par une fonction, on a affaire à une fonction aléatoire. Une fonction aléatoire est donnée donc par un schéma probabiliste arbitraire décrivant des expériences dont les résultats sont des fonctions aléatoires. Cette définition fera l'objet d'un examen plus détaillé au chapitre IV. La définition de la fonction aléatoire proposée dans ce paragraphe sera appelée par convention définition de la fonction aléatoire au sens large.

Jusqu'ici il n'a été question que d'une fonction aléatoire. Or l'on a souvent à résoudre des problèmes qui font intervenir plusieurs fonctions aléatoires. Pour que l'on ait la possibilité d'effectuer des opérations mathématiques sur ces fonctions, il ne suffit pas que chacune d'elles soit donnée séparément. Dans ce cas il est plus simple de remplacer la suite de fonctions $\xi_1(\theta), \xi_2(\theta), \ldots, \xi_m(\theta)$ par une seule fonction vectorielle $\zeta(\theta)$ dont les composantes seraient les fonctions aléatoires précitées. La définition précédente reste pratiquement en vigueur. La répartition de la suite de variables aléatoires (1) sera la fonction de répartition conjointe de la suite de vecteurs $\zeta(\theta_1), \zeta(\theta_2), \ldots, \zeta(\theta_n)$, c'est-à-dire la fonction de nm variables

$$F_{\theta_1, \theta_2, \ldots, \theta_n}(x_{11}, x_{12}, \ldots, x_{nm}) = \\ = \mathsf{P}\{\xi_1(\theta_1) < x_{11},\ \xi_2(\theta_1) < x_{12}, \ldots, \xi_n(\theta_m) < x_{mn}\}.$$

Dans la suite Θ représentera, le plus souvent, l'ensemble des réels et la variable Θ, le temps t. Dans ce cas l'ensemble Θ sera noté $\mathcal{T}$ et désignera un intervalle fini ou infini (fermé, ouvert ou semi-ouvert). On étudie également les cas où $\mathcal{T}$ est composé de tous les entiers non négatifs ou de tous les entiers réels. On a alors affaire à la suite de variables (vecteurs) aléatoires $\zeta(k)$ ($k = 0, 1, 2, \ldots$, ou $k = 0, \pm 1, \pm 2, \ldots$) et le processus est dit *processus aléatoire à temps discret* ou *suite aléatoire*. Les processus à temps discret tiennent une place importante dans la théorie générale des processus

aléatoires. D'abord, ils interviennent dans de nombreux problèmes à temps discret. Ensuite, leur étude peut être entreprise par des méthodes relativement plus simples. Enfin, ils sont susceptibles d'approcher des processus à temps continu.

Dans ce paragraphe nous nous limiterons essentiellement aux fonctions aléatoires à valeurs réelles. Le passage au cas vectoriel n'apporte que des complications d'ordre technique.

Les fonctions de répartition finidimensionnelles définissent de façon univoque une famille de mesures $q_{\theta_1, \theta_2, \ldots, \theta_n}(B)$, $n = 1, 2, \ldots$, $\theta_k \in \Theta$, où B est l'ensemble borélien de $\mathscr{R}^n$. La quantité $q_{\theta_1, \ldots, \theta_n}(B)$ définit la probabilité que dans une expérience la mesure des variables $\xi(\theta_1)$, $\xi(\theta_2)$, ..., $\xi(\theta_n)$ donne une suite de B:

$$q_{\theta_1, \ldots, \theta_n}(B) = \mathsf{P}\{(\xi(\theta_1), \ldots, \xi(\theta_n)) \in B\}.$$

La mesure $q_{\theta_1, \ldots, \theta_n}(B)$ est appelée *répartition de la suite* $\xi(\theta_1), \ldots, \xi(\theta_n)$.

Les expressions des fonctions de répartition finidimensionnelles sont souvent compliquées et d'un maniement peu commode. Aussi préfère-t-on les définir par leurs densités ou leurs fonctions caractéristiques.

Si $f_{\theta_1, \ldots, \theta_n}(x_1, \ldots, x_n)$ est la densité de probabilité correspondant à la fonction de répartition $F_{\theta_1, \ldots, \theta_n}(x_1, \ldots, x_n)$, on a

$$F_{\theta_1, \ldots, \theta_n}(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{\theta_1, \ldots, \theta_n}(y_1, \ldots, y_n)\, dy_1 \ldots dy_n.$$

D'où il suit en particulier

$$f_{\theta_1, \ldots, \theta_n}(x_1, \ldots, x_n) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\theta_1, \ldots, \theta_n, \theta_{n+1}, \ldots, \theta_{n+p}} \times (x_1, \ldots, x_n, y_1, \ldots, y_p)\, dy_1, \ldots, dy_p. \quad (5)$$

Cette formule peut être regardée comme l'équivalent de la condition de compatibilité (4). Entre les mesures $q_{\theta_1, \ldots, \theta_n}(B)$ et les densités de probabilité on a la relation

$$q_{\theta_1, \ldots, \theta_n}(B) = \int \cdots \int_B f_{\theta_1, \ldots, \theta_n}(y_1, \ldots, y_n)\, dy_1, \ldots, dy_n.$$

La fonction caractéristique de la répartition de la suite (1) est donnée par

$$\varphi_{\theta_1, \ldots, \theta_n}(u_1, \ldots, u_n) = \mathsf{E} \exp\left\{ i \sum_{k=1}^{n} \xi(\theta_k)\, u_k \right\},$$

où E est le symbole de l'espérance mathématique, $u_1, \ldots, u_n$ des nombres réels. Si existent les densités de probabilité finidimensionnelles, alors

$$\varphi_{\theta_1, \ldots, \theta_n}(u_1, \ldots, u_n) = \int_{\mathcal{R}^n} \cdots \int e^{i \sum_1^n x_k u_k} f_{\theta_1, \ldots, \theta_n}(x_1, \ldots, x_n)\, dx_1 \ldots dx_n, \quad (6)$$

c'est-à-dire la fonction caractéristique est la transformée de Fourier de la densité de probabilité.

A noter que si l'on peut se donner les répartitions au moyen de leurs fonctions caractéristiques, c'est que ces dernières définissent de façon unique les fonctions de répartition. Si par exemple existent les densités de probabilité finidimensionnelles et qu'elles satisfassent certaines conditions analytiques (étudiées en détail dans la théorie de l'intégrale de Fourier), la densité $f_{\theta_1, \ldots, \theta_n}(x_1, \ldots, x_n)$ peut être calculée à l'aide des fonctions caractéristiques avec la formule de Fourier

$$f_{\theta_1, \ldots, \theta_n}(x_1, \ldots, x_n) = \frac{1}{(2\pi)^n} \int_{\mathcal{R}^n} \cdots \int e^{-\sum_{k=1}^n x_k u_k} \varphi_{\theta_1, \ldots, \theta_n}(u_1, \ldots, u_n)\, du_1, \ldots, du_n. \quad (7)$$

Pour plus de détail sur les fonctions caractéristiques on pourra consulter par exemple W. Feller [2] ou P. Hennequin et A. Tortrat [1].

Fonctions de corrélation. La famille de répartitions conjointes (2) décrit entièrement la fonction aléatoire au sens large. Or souvent on peut se contenter d'une caractéristique plus concise des répartitions, qui précise quelques propriétés importantes de la fonction aléatoire. Par ailleurs la solution de nombreux problèmes probabilistes ne dépend que d'un nombre peu important de paramètres caractérisant les répartitions intervenant dans le problème. Les répartitions sont le mieux caractérisées par leurs *moments.* En théorie des fonctions aléatoires ce sont les fonctions moments qui jouent le rôle de moments des répartitions.

DÉFINITION. *On appelle fonctions moments d'une fonction aléatoire* $\xi(\theta)$, $\theta \in \Theta$, *les fonctions*

$$m_{j_1, j_2, \ldots, j_s}(\theta_1, \theta_2, \ldots, \theta_s) = \mathsf{E}\, [\xi(\theta_1)]^{j_1} [\xi(\theta_2)]^{j_2} \ldots [\xi(\theta_s)]^{j_s},$$

$$j_k \geqslant 0 \quad (k = 1, 2, \ldots, s),$$

si existe l'espérance mathématique du second membre pour tous les $\theta_i \in \Theta$, $i = 1, 2, \ldots, s$. *La quantité* $q = j_1 + j_2 + \ldots + j_s$ *est appelée ordre de la fonction moment.*

DÉFINITION. *Une fonction aléatoire* $\xi(\theta)$, $\theta \in \Theta$, *appartient à la classe* $\mathcal{L}_p(\Theta)$ $(\xi(\theta) \in \mathcal{L}_p(\Theta))$ *si* $\mathsf{E}\,|\xi(\theta)|^p < \infty$ *quel que soit* $\theta \in \Theta$.

On remarque aussitôt que si $\xi(\theta) \in \mathcal{L}_p(\Theta)$, les fonctions moments d'ordre q sont finies pour tous les $q \leqslant p$.

En effet, l'inégalité entre la moyenne géométrique et la moyenne arithmétique

$$\prod_{k=1}^{s} a_a^{p_k} \leqslant \sum_{k=1}^{s} p_k a_k, \quad a_k \geqslant 0, \quad p_k \geqslant 0, \quad \sum_{k=1}^{s} p_k = 1,$$

entraîne

$$\prod_{k=1}^{s} |\xi(\theta_k)|^{j_k} = \prod_{k=1}^{s} |\xi(\theta_k)|^{q\frac{j_k}{q}} \leqslant \sum_{k=1}^{s} \frac{j_k}{q} |\xi(\theta_k)|^q, \quad q = \sum_{k=1}^{s} j_k.$$

L'inégalité de Jensen $\mathsf{E}f(\xi) \leqslant f(\mathsf{E}\xi)$, qui est valable pour toute fonction convexe continue, entraîne

$$\mathsf{E}\prod_{k=1}^{s} |\xi(\theta_k)|^{j_k} \leqslant \mathsf{E}\sum_{k=1}^{s} \frac{j_k}{q} |\xi(\theta_k)|^q \leqslant \sum_{k=1}^{s} \frac{j_k}{q} (\mathsf{E}\,|\xi(\theta_k)|^p)^{\frac{q}{p}}.$$

De là suit la proposition énoncée.

Si les fonctions caractéristiques de répartitions finidimensionnelles sont connues, les fonctions moments à indices entiers peuvent être déduites par dérivation. En effet, si $\xi(\theta) \in \mathcal{L}_p(\Theta)$, alors

$$m_{j_1, \ldots, j_s}(\theta_1, \ldots, \theta_s) = (-i)^q \left. \frac{\partial^q \varphi_{\theta_1, \ldots, \theta_s}(u_1, \ldots, u_s)}{\partial u_1^{j_1}, \ldots, \partial u_s^{j_s}} \right|_{u_1 = \ldots = u_s = 0}$$

avec $q \leqslant p$ $(q = j_1 + j_2 + \ldots + j_s)$. Cette formule se démontre par dérivation par rapport à $u_1, \ldots, u_s$ de l'expression $\varphi_{\theta_1, \ldots, \theta_s}(u_1, \ldots, u_s) = \mathsf{E}e^{i\sum_1^s u_k \xi(\theta_k)}$ sous le symbole de l'espérance mathématique. On fait souvent appel à la réciproque mais celle-ci n'est pas toujours vraie. Elle l'est cependant pour les moments d'ordre pair.

Outre les fonctions moments, on considère souvent les fonctions moments centrées

$$\overline{m}_{j_1, \ldots, j_s}(\theta_1, \ldots, \theta_s) = \\ = \mathsf{E}\left([\xi(\theta_1) - m_1(\theta_1)]^{j_1} [\xi(\theta_2) - m_1(\theta_2)]^{j_2} \ldots [\xi(\theta_s) - m_1(\theta_s)]^{j_s}\right) \quad (8)$$

qui sont les fonctions moments de la fonction aléatoire centrée $\xi_1(\theta) = \xi(\theta) - m_1(\theta)$ dont l'espérance mathématique est nulle quel que soit $\theta \in \Theta$.

Les plus importantes de ces fonctions sont les fonctions des deux premiers ordres :

$$m(\theta) = m_1(\theta) = \mathsf{E}\xi(\theta), \tag{9}$$

$$R(\theta_1, \theta_2) = m_{11}(\theta_1, \theta_2) = \mathsf{E}[\xi(\theta_1) - m(\theta_1)][\xi(\theta_2) - m(\theta_2)]. \tag{10}$$

La fonction $m(\theta)$ est appelée *valeur moyenne*, la fonction $R(\theta_1, \theta_2)$ *fonction de corrélation*. Si $\theta_1 = \theta_2 = \theta$, la fonction de corrélation n'est autre que la *variance* $\sigma^2(\theta)$ de $\xi(\theta)$: $R(\theta, \theta) = \sigma^2(\theta)$.

La quantité

$$r(\theta_1, \theta_2) = \frac{R(\theta_1, \theta_2)}{\sigma(\theta_1)\sigma(\theta_2)} = \frac{R(\theta_1, \theta_2)}{\sqrt{R(\theta_1, \theta_1) R(\theta_2, \theta_2)}}$$

est appelée *coefficient de corrélation* du couple de variables aléatoires $\xi(\theta_1)$ et $\xi(\theta_2)$.

Si $\xi(\theta_1)$ et $\xi(\theta_2)$ sont indépendantes, le coefficient de corrélation est nul. En général la réciproque n'est pas vraie. Cependant dans le cas particulier important où la répartition conjointe des variables aléatoires $\xi(\theta_1)$ et $\xi(\theta_2)$ est normale, la nullité du coefficient de corrélation ou, ce qui revient au même, la nullité de la fonction de corrélation $R(\theta_1, \theta_2)$ entraîne l'indépendance de ces variables. Deux variables aléatoires ξ et η à moments d'ordre deux finis et telles que

$$R_{\xi,\eta} = \mathsf{E}[(\xi - \mathsf{E}\xi)(\eta - \mathsf{E}\eta)] = 0$$

sont dites *non corrélées*.

D'une façon générale, le coefficient de corrélation d'un couple de variables aléatoires est la mesure de la relation linéaire qui les lie, c'est-à-dire le coefficient de corrélation montre avec quelle précision l'une des variables peut être exprimée linéairement en fonction de l'autre.

On a souvent affaire à des fonctions aléatoires à valeurs complexes. On peut les mettre sous la forme $\zeta(\theta) = \xi(\theta) + i\eta(\theta)$ et les considérer comme des fonctions aléatoires vectorielles bidimensionnelles.

S'agissant d'une fonction à valeurs complexes, la relation $\zeta(\theta) \in \in \mathscr{L}_2(\Theta)$ signifie que $\mathsf{E}\,|\zeta(\theta)|^2 < \infty$, $\theta \in \Theta$, c'est-à-dire que $\xi(\theta) \in \mathscr{L}_2(\Theta)$ et $\eta(\theta) \in \mathscr{L}_2(\Theta)$.

La fonction de corrélation d'une fonction aléatoire complexe est définie par

$$R(\theta_1, \theta_2) = \mathsf{E}([\zeta(\theta_1) - \mathsf{E}\zeta(\theta_1)]\,\overline{[\zeta(\theta_2) - \mathsf{E}\zeta(\theta_2)]}),$$

où la barre indique que l'on considère l'expression conjuguée complexe.

Signalons quelques propriétés des fonctions de corrélation :

1) $R(\theta, \theta) \geqslant 0$, l'égalité étant réalisée si et seulement si $\zeta(\theta)$ est constant presque sûrement;

2) $$R(\theta_1, \theta_2) = \overline{R(\theta_2, \theta_1)}; \tag{11}$$

3) $$|R(\theta_1, \theta_2)|^2 \leqslant R(\theta_1, \theta_1) R(\theta_2, \theta_2); \tag{12}$$

4) quels que soient n, $\theta_1, \theta_2, \ldots, \theta_n$ et les complexes $\lambda_1, \lambda_2, \ldots$ $\ldots, \lambda_n$, on a

$$\sum_{j,k=1}^{n} R(\theta_j, \theta_k) \lambda_j \bar{\lambda}_k \geqslant 0. \tag{13}$$

Les deux premières assertions sont évidentes; la troisième est une conséquence de l'inégalité de Cauchy-Bouniakovski

$$(\mathsf{E}\,|\xi\eta|)^2 \leqslant \mathsf{E}\,|\xi|^2 \mathsf{E}\,|\eta|^2.$$

Pour prouver la quatrième on remarquera que

$$\sum_{j,k=1}^{n} R(\theta_j, \theta_k) \lambda_j \bar{\lambda}_k = \mathsf{E} \sum_{j,k=1}^{n} \zeta(\theta_j) \overline{\zeta(\theta_k)} \lambda_j \bar{\lambda}_k = \mathsf{E} \left| \sum_{j=1}^{n} \zeta(\theta_j) \lambda_j \right|^2 \geqslant 0.$$

On notera que les propriétés 1), 2) et 3) découlent de la propriété 4).

Toute fonction $R(\theta_1, \theta_2)$ définie sur Θ et vérifiant la propriété 4) est appelée *noyau défini positif sur* Θ.

La fonction de corrélation mutuelle sert à caractériser le degré de dépendance linéaire de deux fonctions aléatoires $\zeta_1(\theta)$ et $\zeta_2(\theta)$ $(\in \mathscr{L}_2(\Theta))$.

DÉFINITION. *On appelle fonction de corrélation mutuelle de deux fonctions aléatoires* $\zeta_1(\theta)$ *et* $\zeta_2(\theta)$ $(\zeta_1(\theta), \zeta_2(\theta) \in \mathscr{L}_2(\Theta))$ *la fonction*

$$R_{\zeta_1\zeta_2}(\theta_1, \theta_2) = \mathsf{E}\,[\zeta_1(\theta_1) - \mathsf{E}\zeta_1(\theta_1)]\,\overline{[\zeta_2(\theta_2) - \mathsf{E}\zeta_2(\theta_2)]}.$$

Soit donnée la suite de fonctions aléatoires à valeurs complexes:

$$\zeta_1(\theta), \zeta_2(\theta), \ldots, \zeta_r(\theta), \quad \zeta_i(\theta) \in \mathscr{L}_2(\Theta), \quad i = 1, 2, \ldots, r.$$

On conviendra de la considérer comme une fonction aléatoire complexe r-dimensionnelle

$$\zeta(\theta) = \{\zeta_1(\theta), \zeta_2(\theta), \ldots, \zeta_r(\theta)\}, \quad \theta \in \Theta.$$

Etant donné deux vecteurs ξ et η,

$$\xi = (\xi_1, \xi_2, \ldots, \xi_r),$$
$$\eta = (\eta_1, \eta_2, \ldots, \eta_r),$$

on désignera par $\xi\eta^*$ la matrice

$$\xi\eta^* = \begin{pmatrix} \xi_1\bar{\eta}_1 & \xi_1\bar{\eta}_2 & \cdots & \xi_1\bar{\eta}_r \\ \xi_2\bar{\eta}_1 & \xi_2\bar{\eta}_2 & \cdots & \xi_2\bar{\eta}_r \\ \cdot & \cdot & \cdot & \cdot \\ \xi_r\bar{\eta}_1 & \xi_r\bar{\eta}_2 & \cdots & \xi_r\bar{\eta}_r \end{pmatrix} = (\xi_i\bar{\eta}_j)_{i,j=1,\ldots,r}.$$

Posons

$$m(\theta) = \mathsf{E}\,\zeta(\theta) = \{\mathsf{E}\zeta_1(\theta),\ \mathsf{E}\zeta_2(\theta),\ \ldots,\ \mathsf{E}\zeta_r(\theta)\},$$

$$\begin{aligned} R(\theta_1, \theta_2) &= (R_{ij}(\theta_1, \theta_2))_{i,j=1,\ldots,r} = \\ &= \mathsf{E}\,([\zeta(\theta_1) - m(\theta_1)]\,[\zeta(\theta_2) - m(\theta_2)]^*) = \\ &= (\mathsf{E}\{[\zeta_i(\theta_1) - m_i(\theta_1)]\,\overline{[\zeta_j(\theta_2) - m_j(\theta_2)]}\})_{i,j=1,\ldots,r}. \end{aligned}$$

La fonction $m(\theta)$ est une fonction vectorielle complexe r-dimensionnelle. On l'appelle *valeur moyenne* de la fonction aléatoire vectorielle $\zeta(\theta)$. La matrice $R(\theta_1, \theta_2)$ s'appelle *matrice de corrélation de* $\zeta(\theta)$.

Aux propriétés 1) à 4) des fonctions de corrélation correspondent les propriétés suivantes de la matrice de corrélation d'une fonction aléatoire:

1) $R(\theta, \theta)$ est une matrice définie non négative

$$\sum_{j,k=1}^{r} R_{jk}(\theta, \theta)\,\lambda_j\bar{\lambda}_k = \mathsf{E}\,\Big|\sum_{j=1}^{r} \lambda_j\zeta_j(\theta)\Big|^2 \geqslant 0; \tag{14}$$

2) $$R(\theta_1, \theta_2)^* = R(\theta_2, \theta_1), \tag{15}$$

où * désigne la matrice conjuguée complexe;

3) $$|R_{jk}(\theta_1, \theta_2)|^2 \leqslant R_{jj}(\theta_1, \theta_1)\,R_{kk}(\theta_2, \theta_2),\quad j, k = 1, \ldots, r; \tag{16}$$

4) quels que soient n, $\theta_1, \ldots, \theta_n$ et la suite de vecteurs complexes $A_1, A_2, \ldots, A_n$, on a

$$\sum_{j,k=1}^{n} (R(\theta_j, \theta_k)\,A_k, A_j) \geqslant 0. \tag{17}$$

La dernière condition équivaut à la suivante:

4′) quelle que soit la suite de matrices $\Lambda_1, \ldots, \Lambda_n$, la matrice

$$\sum_{j,k=1}^{n} \Lambda_j R(\theta_j, \theta_k)\,\Lambda_k^*$$

est définie non négative.

Les propriétés 1) et 2) sont évidentes. Pour démontrer la propriété 3) on fera appel à l'inégalité de Cauchy-Bouniakovski relative à l'espérance mathématique:

$$\begin{aligned} |R_{jk}(\theta_1, \theta_2)|^2 = |\,\mathsf{E}\,[\zeta_j(\theta_1) - m_j(\theta_1))\,\overline{(\zeta_k(\theta_2) - m_k(\theta_2))}]\,|^2 \leqslant \\ \leqslant R_{jj}(\theta_1, \theta_1)\,R_{kk}(\theta_2, \theta_2). \end{aligned}$$

Pour prouver la propriété 4) on posera $A_k = (a_{k1}, \ldots, a_{kr})$. Il vient alors

$$\begin{aligned} \sum_{j,k=1}^{n} (R(\theta_j, \theta_k)\,A_k, A_j) &= \sum_{j,k=1}^{n} \sum_{p,q=1}^{r} R_{pq}(\theta_j, \theta_k)\,a_{kq}\bar{a}_{jp} = \\ &= \mathsf{E}\,\Big|\sum_{j=1}^{n} \sum_{p=1}^{r} (\zeta_p(\theta_j) - m_p(\theta_j))\,\bar{a}_{jp}\Big|^2 \geqslant 0. \end{aligned}$$

Les notions de continuité, mesurabilité et autres ne s'appliquent pas directement aux fonctions aléatoires au sens large. En principe il est possible de leur trouver des notions équivalentes en termes de répartitions finidimensionnelles de la fonction aléatoire. Cependant l'approche que nous proposons au chapitre IV est actuellement la plus répandue. Rappelons une notion fondée sur la répartition des couples $\xi(\theta_1)$, $\xi(\theta_2)$.

Soient $\xi(\theta)$, fonction vectorielle à valeurs dans $\mathcal{R}^d$, Θ, un espace métrique muni de la métrique $r(\theta_1, \theta_2)$.

DÉFINITION. On dit qu'*une fonction aléatoire* $\xi(\theta)$, $\theta \in \Theta$, *est stochastiquement continue en un point* θ_0 *si quel que soit* $\varepsilon > 0$

$$\mathsf{P}\{|\xi(\theta_0) - \xi(\theta)| > \varepsilon\} \to 0 \textit{ lorsque } r(\theta_0, \theta) \to 0.$$

Si $\xi(\theta)$ est stochastiquement continue en tout point d'un ensemble $B \subset \Theta$, on dit qu'elle est *stochastiquement continue sur* B.

La continuité stochastique d'une fonction aléatoire sur B n'implique pas nécessairement la continuité de ses réalisations sur B. Quelques exemples simples le prouvent (cf. § 3).

DÉFINITION. *Une fonction aléatoire* $\xi(\theta)$ *est dite stochastiquement bornée sur* B *si*

$$\sup_{\theta \in B} \mathsf{P}\{|\xi(\theta)| > N\} \to 0$$

lorsque $N \to \infty$.

THÉORÈME 1. *Une fonction aléatoire* $\xi(\theta)$ *stochastiquement continue sur un compact* Θ *est stochastiquement bornée sur ce compact.*

Démonstration. Soit ε un nombre positif arbitrairement petit. A tout point $\theta \in \Theta$ associons une sphère S_θ centrée en θ et telle que

$$\mathsf{P}\{|\xi(\theta) - \xi(\theta')| > 1\} < \frac{\varepsilon}{2} \ \forall \theta' \in S_\theta.$$

Parmi l'ensemble de sphères S_θ choisissons un recouvrement fini $S_{\theta_1}, \ldots, S_{\theta_n}$ de l'ensemble Θ. Soit a le rayon de la plus grande des sphères $S_{\theta_1}, \ldots, S_{\theta_n}$. Il vient

$$|\xi(\theta)| \leqslant |\xi(\theta) - \xi(\theta_j)| + \max |\xi(\theta_j)|,$$

θ_j étant le centre de la sphère S_{θ_j} où tombe le point θ. Donc pour $N > 2$ on a

$$\mathsf{P}\{|\xi(\theta)| > N\} \leqslant \mathsf{P}\left\{|\xi(\theta) - \xi(\theta_j)| > \frac{N}{2}\right\} + \\ + \mathsf{P}\left\{\max |\xi(\theta_j)| > \frac{N}{2}\right\} \leqslant \frac{\varepsilon}{2} + \mathsf{P}\left\{\max_{1 \leqslant j \leqslant n} |\xi(\theta_j)| > \frac{N}{2}\right\}.$$

La quantité $\max_j |\xi(\theta_j)|$ est finie presque sûrement. Par conséquent, si N est suffisamment grand, c'est-à-dire $N \geqslant N_0(\varepsilon)$, on a

$$\mathsf{P}\left\{\max_{1\leqslant j\leqslant n} |\xi(\theta_j)| > \frac{N}{2}\right\} < \frac{\varepsilon}{2} \text{ et } \mathsf{P}\{\xi(\theta)| > N\} < \varepsilon. \blacksquare$$

DÉFINITION. *Une fonction aléatoire* $\xi(\theta)$ *est uniformément stochastiquement continue sur* Θ *si quel que soit* $\varepsilon > 0$ *on peut exhiber un* $\delta > 0$ *tel que*

$$\mathsf{P}\{|\xi(\theta) - \xi(\theta')| > \varepsilon\} < \varepsilon$$

pour tous les θ *et* θ' *tels que* $r(\theta, \theta') < \delta$.

THÉORÈME 2. *Une fonction aléatoire* $\xi(\theta)$ *stochastiquement continue sur un compact* Θ *est uniformément stochastiquement continue sur* Θ.

Admettons le contraire. Il existerait un $\varepsilon > 0$ et pour tout n un couple de points θ_n et θ'_n pour lesquels $r(\theta_n, \theta'_n) < \frac{1}{n}$ et $\mathsf{P}\{|\xi(\theta_n) - \xi(\theta'_n)| > \varepsilon\} \geqslant \varepsilon$. L'ensemble Θ étant compact, on peut extraire une suite partielle d'indices n_k tels que θ_{n_k} et θ'_{n_k} convergent vers une limite θ_0. Alors

$$\varepsilon \leqslant \mathsf{P}\{|\xi(\theta_{n_k}) - \xi(\theta'_{n_k})| > \varepsilon\} \leqslant \mathsf{P}\left\{|\xi(\theta_{n_k}) - \xi(\theta_0)| > \frac{\varepsilon}{2}\right\} + \\ + \mathsf{P}\left\{|\xi(\theta_0) - \xi(\theta'_{n_k})| > \frac{\varepsilon}{2}\right\}.$$

La continuité stochastique de $\xi(\theta)$ entraîne que le second membre de la dernière inégalité tend vers 0 avec k. Cette contradiction démontre le théorème.

§ 2. Fonctions aléatoires gaussiennes

Les fonctions aléatoires admettant des répartitions gaussiennes (normales) finidimensionnelles sont très importantes dans les applications. Définissons et étudions les principales propriétés de la répartition gaussienne à n dimensions.

DÉFINITION. *Un vecteur aléatoire* $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ *admet une répartition gaussienne si la fonction caractéristique de la répartition est de la forme*

$$\varphi(u) = \mathsf{E}e^{i(u,\xi)} = e^{i(m,u) - \frac{1}{2}(Ru,u)}, \tag{1}$$

où $m = (m_1, m_2, \ldots, m_n)$, $u = (u_1, u_2, \ldots, u_n)$ *sont des vecteurs,* $R = (r_{ik})$, $i, k = 1, 2, \ldots, n$, *est une matrice symétrique réelle définie non négative. Par* (α, β) *on désigne le produit scalaire des vecteurs* α *et* β, *de sorte que*

$$(m, u) = \sum_{k=1}^{n} m_k u_k, \quad (Ru, u) = \sum_{j,k=1}^{n} r_{jk} u_j u_k.$$

Le théorème suivant est une justification formelle de la définition précédente.

THÉORÈME 1. *Pour que la fonction*

$$\psi(u)=\exp\left\{i(m,\ u)-\frac{1}{2}(Ru,\ u)\right\}$$

soit fonction caractéristique de la répartition d'un vecteur aléatoire ξ *à* n *dimensions, il est nécessaire et suffisant que la matrice réelle* R *soit définie non négative, symétrique et de rang égal à la dimension du sous-espace contenant la répartition du vecteur* ξ.

D é m o n s t r a t i o n. a) Condition nécessaire. Supposons que la fonction caractéristique $\varphi(u)$ d'un vecteur aléatoire est définie par (1). En dérivant par rapport à u_j et ensuite par rapport à u_k et en faisant $u=0$, on constate que la répartition possède des moments finis et

$$\left.\frac{\partial\varphi}{\partial u_j}\right|_{u=0}=i\mathsf{E}\xi_j=im_j, \tag{2}$$

$$\left.\frac{\partial^2\varphi}{\partial u_j\partial u_k}\right|_{u=0}=-\mathsf{E}\xi_j\xi_k=-m_jm_k-r_{jk}. \tag{3}$$

Ces formules entraînent que la matrice R est réelle, symétrique et définie non négative:

$$(Ru,\ u)=\mathsf{E}\Big(\sum_{j=1}^{k}(\xi_j-m_j)\,u_j\Big)^2=\operatorname{Var}(\xi,\ u)\geqslant 0. \tag{4}$$

Si le rang de la matrice R est égal à $r\,(\leqslant n)$, par un changement convenable des variables $u_j=\sum_{k=1}^{n}\alpha_{jk}v_k$ on peut la réduire à ses axes principaux:

$$(Ru,\ u)=\sum_{k=1}^{r}\lambda_k v_k^2=\mathsf{E}\,\Big[\sum_{k=1}^{n}\sum_{j=1}^{n}(\xi_j-m_j)\,\alpha_{jk}v_k\Big]^2.$$

Donc $\sum_{j=1}^{n}(\xi_j-m_j)\,\alpha_{jk}=0$ pour $k=r+1,\ \ldots,\ n$ presque sûrement. Ces relations montrent qu'entre les composantes du vecteur ξ il existe presque sûrement $n-r$ relations indépendantes, donc sa répartition est concentrée dans un hyperplan de dimension r défini par les équations

$$\sum_{j=1}^{n}(x_j-m_j)\,\alpha_{jk}=0,\quad k=r+1,\ \ldots,\ n.$$

b) Condition suffisante. Supposons que R est une matrice symétrique définie positive. La fonction $\psi(u)=\exp\left\{i\,(m,\,u)-\frac{1}{2}\,(Ru,\,u)\right\}$

est absolument intégrable et dérivable. On peut donc lui appliquer la formule intégrale de Fourier:

$$\begin{aligned} \psi(u) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x)\, e^{i(x,u)}\, dx_1 \ldots dx_n, \\ f(x) &= \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \psi(u)\, e^{-i(u,x)}\, du_1 \ldots du_n. \end{aligned} \tag{5}$$

Les intégrales du second membre sont n-uples.

Soit C la matrice orthogonale qui réduit R à la forme diagonale, c'est-à-dire $C^*RC = D$, où $D = (\lambda_i \delta_{ik})$, $i, k = 1, \ldots, n$; $\lambda_i > 0$, C^* est la matrice conjuguée de C. La matrice C étant réelle et orthogonale, C^* est confondue avec la transposée et l'inverse de C, c'est-à-dire $C^* = C' = C^{-1}$. Faisons le changement de variables $u = Cv$ ou $v = C^*u$, où $v = (v_1, v_2, \ldots, v_n)$. L'élément de volume étant invariant par la transformation orthogonale, on a

$$f(x) = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left\{-i(x-m,\ Cv) - \frac{1}{2}(RCv,\ Cv)\right\} dv_1 \ldots dv_n.$$

Par ailleurs

$$(RCv,\ Cv) = (C^*RCv,\ v) = \sum_{k=1}^{n} \lambda_k v_k^2,$$

$$(x-m,\ Cv) = (C^*(x-m), v) = \sum_{k=1}^{n} x_k^* v_k,$$

où x_k^* est la k-ième composante du vecteur $x^* = C^*(x-m)$. Donc

$$\begin{aligned} f(x) &= \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left\{-i \sum_{k=1}^{n} x_k^* v_k - \frac{1}{2} \sum_{k=1}^{n} \lambda_k v_k^2\right\} dv_1 \ldots dv_n = \\ &= \prod_{k=1}^{n} \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left\{-i x_k^* v_k - \frac{1}{2} \lambda_k v_k^2\right\} dv_k = \\ &= \prod_{k=1}^{n} \frac{1}{\sqrt{2\pi\lambda_k}} e^{-x_k^{*2}/2\lambda_k} = (2\pi)^{-\frac{n}{2}} \left(\prod_{k=1}^{n} \lambda_k\right)^{-\frac{1}{2}} e^{-\frac{1}{2}(D^{-1}x^*,\ x^*)}. \end{aligned}$$

D'autre part $\prod_{k=1}^{n} \lambda_k = \Delta$, où Δ est le déterminant de la matrice R,

$$(D^{-1}x^*, x^*) = (D^{-1}C^* (x - m), C^* (x - m)) =$$
$$= (CD^{-1}C^* (x - m), (x - m)) =$$
$$= ((CDC^*)^{-1} (x - m), (x - m)) = (R^{-1} (x - m), (x - m)),$$

R^{-1} étant l'inverse de la matrice R. On obtient en définitive

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \Delta}} \exp\left\{-\frac{1}{2}(R^{-1}(x-m), (x-m))\right\} =$$
$$= \frac{1}{\sqrt{(2\pi)^n \Delta}} \exp\left\{-\frac{1}{2} \sum_{j,k=1}^{n} \frac{\Delta_{kj}(x_j - m_j)(x_k - m_k)}{\Delta}\right\}, \quad (6)$$

où Δ_{kj} sont les cofacteurs des éléments de R. De (6) on déduit que $f(x) > 0$ et de (5) que $\int_{\infty}^{\infty} f(x)\, dx = \psi(0) = 1$. Donc la fonction $f(x)$ peut être considérée comme une densité de probabilité n-dimensionnelle et $\psi(u)$ comme sa fonction caractéristique.

Généralisons. Soient R de rang r $(r < n)$ et C, transformation orthogonale diagonalisant R, c'est-à-dire $C^*RC = D_r$, où D_r est une matrice diagonale dont les éléments diagonaux $\lambda_k = 0$ pour $k = r + 1, \ldots, n$ et $\lambda_k > 0$ pour $k = 1, 2, \ldots, r$. Soit $\lambda_j^\varepsilon = \lambda_j$ pour $j = 1, 2, \ldots, r$; $\lambda_j^\varepsilon = \varepsilon$ pour $j = r + 1, \ldots, n$. Alors $R_\varepsilon = CD_\varepsilon C^*$ est une matrice définie positive (D_ε est une matrice d'éléments $\lambda_j^\varepsilon \delta_{jk}$) et

$$\varphi_\varepsilon(u) = \exp\left\{i(m, u) - \frac{1}{2}(R_\varepsilon u, u)\right\}$$

est fonction caractéristique d'une répartition. Lorsque $\varepsilon \to 0$, la fonction $\varphi_\varepsilon(u)$ tend uniformément vers $\varphi(u)$, donc $\varphi(u)$ sera également fonction caractéristique d'une répartition. On a vu précédemment que cette répartition est concentrée dans un hyperplan de r dimensions, donc elle ne possède pas de densité. Une telle répartition s'appelle *répartition gaussienne impropre*. ■

COROLLAIRE 1. *Dans l'expression* (1) *de la fonction caractéristique de la répartition gaussienne*, $m = (m_1, m_2, \ldots, m_n)$ *est le vecteur de l'espérance mathématique*, R *la matrice de corrélation*:

$$m = \mathsf{E}\xi, \quad r_{jk} = \mathsf{E}[(\xi_j - m_j)(\xi_k - m_k)].$$

Ce corollaire découle immédiatement des formules (2) et (3).

COROLLAIRE 2. *Si la matrice de corrélation R d'un vecteur aléatoire gaussien ξ est non dégénérée, il existe une densité de probabilité n-dimensionnelle $f(x)$ définie par* (6).

COROLLAIRE 3. *La répartition conjointe d'un groupe quelconque de composantes d'un vecteur aléatoire gaussien est gaussienne.*

THÉORÈME 2. *Si un vecteur aléatoire $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ est à répartition gaussienne et les vecteurs aléatoires $\xi' = (\xi_1, \ldots, \xi_r)$, $\xi'' = (\xi_{r+1}, \ldots, \xi_n)$ $(r < n)$ ne sont pas corrélés, alors les vecteurs ξ' et ξ'' sont indépendants.*

Démonstration. La non-corrélation des vecteurs ξ' et ξ'' entraîne

$$\mathsf{E}\xi_i\xi_j - \mathsf{E}\xi_i\mathsf{E}\xi_j = 0, \quad i = 1, \ldots, r, \quad j = r+1, \ldots, n,$$

donc

$$\varphi(u) = \exp\Big\{i(m', u') + i(m'', u'') - \\ -\frac{1}{2}\sum_{j,k=1}^{r} r_{jk}u_ju_k - \frac{1}{2}\sum_{j,k=r+1}^{n} r_{jk}u_ju_k\Big\},$$

où

$$m' = (m_1, m_2, \ldots, m_r), \qquad m'' = (m_{r+1}, \ldots, m_n),$$
$$u' = (u_1, u_2, \ldots, u_r), \qquad u'' = (u_{r+1}, \ldots, u_n).$$

La formule précédente s'écrit encore

$$\varphi(u) = \mathsf{E}e^{i(u', \xi') + i(u'', \xi'')} = \mathsf{E}e^{i(u', \xi')}\mathsf{E}e^{i(u'', \xi'')} = \varphi'(u')\,\varphi''(u''),$$

où $\varphi'(u')$ et $\varphi''(u'')$ sont les fonctions caractéristiques des vecteurs ξ' et ξ''. Cette relation prouve l'indépendance de ξ' et ξ''.

Soit $A = \| a_{jk} \|$ $(j = 1, \ldots, h,\ k = 1, \ldots, n)$ une matrice rectangulaire et $\eta = A\xi$, c'est-à-dire

$$\eta = (\eta_1, \ldots, \eta_h), \qquad \eta_j = \sum_{k=1}^{n} a_{jk}\xi_k, \qquad j = 1, \ldots, h.$$

Le vecteur η est associé au vecteur ξ par une application linéaire.

THÉORÈME 3. *L'image d'une répartition gaussienne par une application linéaire de vecteurs gaussiens est une répartition gaussienne.*

Démonstration. Soit $\varphi_\eta(t_1, \ldots, t_n)$ la fonction caractéristique du vecteur η. On a

$$\varphi_\eta(u_1, \ldots, u_h) = \mathsf{E}\exp\Big\{i\sum_{j=1}^{h} u_j\eta_j\Big\} = \\ = \mathsf{E}\exp\Big\{i\sum_{k=1}^{n}\Big(\sum_{j=1}^{h} u_ja_{jk}\Big)\xi_k\Big\} = \exp\Big\{i(u, Am) - \frac{1}{2}(ARA'u, u)\Big\}, \tag{7}$$

c'est-à-dire η admet une répartition gaussienne d'espérance mathématique Am et de matrice de variance $R_\eta = ARA'$. ■

Etablissons quelques formules souvent usitées, relatives aux répartitions gaussiennes.

Soient $\xi = (\xi_1, \ldots, \xi_n)$ et $\eta = (\eta_1, \ldots, \eta_n)$ deux vecteurs admettant une répartition conjointe gaussienne. Déterminons la répartition conditionnelle du vecteur ξ pour η donné. Sans nuire à la généralité on peut supposer que la matrice de corrélation R_{22} du vecteur η est non dégénérée. En effet, dans le cas contraire certaines composantes du vecteur η auraient été des combinaisons linéaires d'autres composantes qu'on aurait pu éliminer et réduire ainsi la dimension du vecteur η.

Soient $m_1 = \mathsf{E}\xi$, $m_2 = \mathsf{E}\eta$, R_{11} la matrice de corrélation du vecteur ξ, R_{12} la matrice de corrélation mutuelle des vecteurs ξ et η :

$$R_{12} = \mathsf{E}\,(\xi - m_1)\,(\eta - m_2)^*.$$

On remarquera que quels que soient la matrice non aléatoire A et les vecteurs ξ et η, on a

$$\mathsf{E}(A\xi, \eta) = \operatorname{Tr} AK,$$

où K est la matrice d'éléments $k_{ij} = \mathsf{E}\xi_i\eta_j$, $\operatorname{Tr} B = \sum_j b_{jj}$ la trace de la matrice B. En outre, $\operatorname{Tr} AB = \operatorname{Tr} BA$, $\forall A$, $\forall B$. Revenons à notre problème et posons

$$\check{\xi} = m_1 + R_{12}R_{22}^{-1}(\eta - m_2). \tag{8}$$

Il est aisé de s'assurer que pour toute matrice A on a

$$\mathsf{E}(A\,(\xi - \check{\xi}),\ \eta - m_2) = 0. \tag{9}$$

En effet,

$$\mathsf{E}\,(A\,(\xi - \check{\xi}),\ \eta - m_2) = \operatorname{Tr} AR_{12} - \operatorname{Tr} AR_{12}R_{22}^{-1}R_{22} = 0.$$

L'égalité (9) signifie qu'aucune composante du vecteur $\xi - \check{\xi}$ n'est corrélée avec une composante du vecteur $\eta - m_2$. Donc ces vecteurs sont indépendants. Par suite

$$\xi = \zeta + \check{\xi} = \zeta + m_1 + R_{12}R_{22}^{-1}(\eta - m_2), \tag{10}$$

où ζ et $\check{\xi}$ sont des vecteurs aléatoires gaussiens et ζ ne dépend pas de η. D'autre part

$$\mathsf{E}\zeta = m_1 - m_1 = 0,$$

$$\mathsf{E}\zeta\zeta^* = \mathsf{E}(\xi - \check{\xi})\,(\xi - \check{\xi})^* =$$

$$= \mathsf{E}\,\{(\xi - m_1)\,(\xi - m_1)^* - (\xi - m_1)\,(\eta - m_2)^*\,R_{22}^{-1}R_{12}^* -$$

$$- R_{12}R_{22}^{-1}(\eta - m_2)\,(\xi - m_1)^* + R_{12}R_{22}^{-1}(\eta - m_2)\,(\eta - m_2)^*\,R_{22}^{-1}R_{12}^*\},$$

ou, puisque $R_{12}^* = R_{21} = \mathsf{E}(\eta - m_2)\,(\xi - m_1)^*$ et la matrice R_{22} est symétrique,

$$\mathsf{E}\zeta\zeta^* = R_{11} - R_{12}R_{22}^{-1}R_{21}.$$

La formule (10) entraîne que la répartition conditionnelle du vecteur ξ pour η donné est gaussienne, de moyenne conditionnelle

$$\mathsf{E}\{\xi \mid \eta\} = \check{\xi} = m_1 + R_{12} R_{22}^{-1} (\eta - m_2) \tag{11}$$

et de matrice de corrélation conditionnelle

$$\mathsf{E}\{(\xi - \check{\xi})(\xi - \check{\xi})^* \mid \eta\} = R_{11} - R_{12} R_{22}^{-1} R_{21}. \tag{12}$$

Signalons un fait important, savoir que la matrice de corrélation conditionnelle du vecteur ξ pour η donné n'est pas aléatoire et, en particulier, ne dépend pas de la valeur du vecteur η.

La proposition suivante est évidente.

THÉORÈME 4. *Soit* $\xi^{(\alpha)}$ $(\alpha = 1, 2, \ldots, r, \ldots)$ *une suite de vecteurs n-dimensionnels admettant une répartition gaussienne de paramètres* $(m^{(\alpha)}, R^{(\alpha)})$. *La suite de répartitions des vecteurs* $\xi^{(\alpha)}$ $(\alpha = 1, 2, \ldots, r)$ *est faiblement convergente vers une répartition limite si et seulement si*

$$m^{(\alpha)} \to m, \quad R^{(\alpha)} \to R. \tag{13}$$

Dans ce cas la répartition limite est également gaussienne et de paramètres (m, R).

Passons maintenant aux fonctions aléatoires. Une fonction aléatoire vectorielle s-dimensionnelle $\xi(\theta) = \{\xi_1(\theta), \ldots, \xi_r(\theta)\}$ est par définition *gaussienne* si la répartition conjointe de toutes les composantes des vecteurs aléatoires

$$\xi(\theta_1), \ \xi(\theta_2), \ \ldots, \ \xi(\theta_n) \tag{14}$$

est gaussienne.

La matrice de corrélation R de la répartition conjointe de la suite de vecteurs aléatoires (14) est de dimension $sn \times sn$ et peut être divisée en blocs carrés de dimension $s \times s$ de la manière suivante :

$$R = \begin{pmatrix} R(\theta_1, \theta_1) & R(\theta_1, \theta_2) & \ldots & R(\theta_1, \theta_n) \\ R(\theta_2, \theta_1) & R(\theta_2, \theta_2) & \ldots & R(\theta_2, \theta_n) \\ \ldots & \ldots & \ldots & \ldots \\ R(\theta_n, \theta_1) & R(\theta_n, \theta_2) & \ldots & R(\theta_n, \theta_n) \end{pmatrix},$$

où $R(\theta_1, \theta_2)$ est la matrice de corrélation de la fonction $\xi(\theta)$. La matrice R est réelle et définie non négative.

La réciproque est évidente. Plus exactement, quelles que soient la fonction vectorielle réelle $m(\theta)$ et la fonction matricielle réelle définie non négative $R(\theta_1, \theta_2)$, $\theta_i \in \Theta$ $(i = 1, 2)$, il existe une fonction aléatoire gaussienne r-dimensionnelle (au sens large) admettant $m(\theta)$ pour vecteur de l'espérance mathématique et $R(\theta_1, \theta_2)$ pour matrice de corrélation.

Les moments d'une fonction aléatoire gaussienne peuvent être obtenus par développement de la fonction caractéristique. On se limitera aux moments centrés. Posons $m(\theta) = 0$. Il s'ensuit

$$\varphi_{\theta_1, \ldots, \theta_s}(u_1, \ldots, u_s) = e^{-\frac{1}{2}(Ru, u)} =$$
$$= 1 - \frac{1}{2}(Ru, u) + \frac{1}{2!2^2}(Ru, u)^2 - \ldots + (-1)^n \frac{1}{2^n n!}(Ru, u)^n + \ldots$$

D'où il vient pour les fonctions moments d'ordre impair

$$m_{j_1, \ldots, j_s}(\theta_1, \ldots, \theta_s) = 0, \text{ si } \sum_{k=1}^{s} j_k = 2n+1,$$

et pour les fonctions moments centrées d'ordre pair

$$m_{j_1, \ldots, j_s}(\theta_1, \ldots, \theta_s) = \frac{\partial^{2n}}{\partial t_1^{j_1} \ldots \partial t_s^{j_s}} \cdot \frac{1}{2^n n!}(Ru, u)^n, \quad \sum_{k=1}^{s} j_k = 2n. \tag{15}$$

Les formules suivantes sont valables pour les fonctions moments d'ordre quatre:

$$m_4(\theta) = 3R^2(\theta, \theta), \quad m_{31}(\theta_1, \theta_2) = 3R(\theta_1, \theta_1) R(\theta_1, \theta_2),$$
$$m_{211}(\theta_1, \theta_2, \theta_3) = R(\theta_1, \theta_1) R(\theta_2, \theta_3) + 2R(\theta_1, \theta_2) R(\theta_1, \theta_3),$$
$$m_{1111}(\theta_1, \theta_2, \theta_3, \theta_4) = R(\theta_1, \theta_2) R(\theta_3, \theta_4) +$$
$$+ R(\theta_1, \theta_3) R(\theta_2, \theta_4) + R(\theta_1, \theta_2) R(\theta_2, \theta_3).$$

Dans le cas général on a la relation

$$m_{j_1, \ldots, j_s}(\theta_1, \ldots, \theta_s) = \sum \prod R(\theta_p, \theta_q), \tag{16}$$

qui se déchiffre comme suit. On écrit la suite de points $\theta_1, \ldots, \theta_s$, le point θ_k étant reproduit j_k fois de suite. On décompose cette suite en couples arbitraires, puis on étend le produit à tous les couples de cette partition et la somme à toutes les partitions (les couples ne sont pas ordonnés). Cette proposition découle immédiatement de la formule (15).

Les fonctions aléatoires gaussiennes jouent un rôle important dans les applications, souvent pour la raison suivante. Lorsque les conditions sont larges, la somme d'un grand nombre de fonctions aléatoires indépendantes et petites est approximativement une fonction aléatoire gaussienne indépendamment de la nature probabiliste de telle ou telle composante. Ceci constitue le théorème de corrélation normale qui est la généralisation du théorème limite central.

DÉFINITION. *Une suite* $\xi_n(\theta)$, $\theta \in \Theta$, $n = 1, 2, \ldots$, *de fonctions aléatoires (au sens large) est faiblement convergente vers* $\xi_0(\theta)$, $\theta \in \Theta$,

si quels que soient s, θ_i ($\theta_i \in \Theta$, $s = 1, 2, \ldots$) *la répartition conjointe des variables aléatoires* ($\xi_n(\theta_1), \ldots, \xi_n(\theta_s)$) *converge faiblement vers la répartition des variables* ($\xi_0(\theta_1), \ldots, \xi_0(\theta_s)$) *lorsque* $n \to \infty$.

THÉORÈME 5. *Soit donnée la suite de sommes de fonctions aléatoires*

$$\eta_n(\theta) = \sum_{k=1}^{m_n} \alpha_{nk}(\theta), \quad \theta \in \Theta, \quad n = 1, 2, \ldots$$

Si

1) *pour* n *fixe les* n *variables aléatoires* $\alpha_{n1}(\theta_1)$, $\alpha_{n2}(\theta_2), \ldots$ $\ldots, \alpha_{nm_n}(\theta_{m_n})$ *sont deux à deux indépendantes quels que soient* $\theta_1, \theta_2, \ldots, \theta_{m_n}$ *et admettent des moments du second ordre tels que*

$$\mathsf{E}\alpha_{nk}(\theta) = 0, \quad \mathsf{E}\alpha_{nk}^2(\theta) = b_{nk}^2(\theta);$$

2) *la fonction de corrélation* $R_n(\theta_1, \theta_2) = \mathsf{E}[\eta_n(\theta_1), \eta_n(\theta_2)]$ *tend vers la limite*

$$\lim_{n\to\infty} R_n(\theta_1, \theta_2) = R(\theta_1, \theta_2);$$

3) *les sommes* $\eta_n(\theta) = \sum_{k=1}^{m_n} \alpha_{nk}(\theta)$ *vérifient la condition de Lindeberg quel que soit* θ: *pour* $\tau > 0$

$$\frac{1}{B_n^2} \sum_{k=1}^{m_n} \int_{|x| > \tau B_n} x^2 \, d\Pi_{nk}(\theta, x) \to 0,$$

où $\Pi_{nk}(\theta, x)$ *est la fonction de répartition de la variable aléatoire* $\alpha_{nk}(\theta)$ *et*

$$B_n^2 = \sum_{k=1}^{m_n} b_{nk}^2(\theta) = R_n(\theta, \theta),$$

alors la fonction aléatoire $\eta_n(\theta)$ *converge faiblement vers une fonction aléatoire admettant une répartition gaussienne d'espérance mathématique nulle et de matrice de corrélation* $R(\theta_1, \theta_2)$ *lorsque* $n \to \infty$.

Ce théorème découle immédiatement du théorème limite central pour les sommes de vecteurs aléatoires indépendants. Les conditions peuvent être affaiblies.

§ 3. Processus à accroissements indépendants

A proprement parler, la théorie des processus aléatoires est née de l'étude des processus à accroissements indépendants. Les recherches ont d'abord porté sur le processus de Wiener (ou mouvement brownien) et ensuite sur des processus plus généraux. Elles ont

consisté à décrire entièrement cette classe de processus et à en étudier les propriétés.

Dans ce paragraphe on se propose de citer des exemples de processus à accroissements indépendants et d'établir l'expression générale des fonctions caractéristiques des répartitions finidimensionnelles de processus stochastiquement continus à accroissements indépendants.

Soit T un intervalle fini $[0, a]$ ou infini $[0, \infty[$.

Définition. *Un processus aléatoire* $\{\xi(t),\ t \in T\}$ *à valeurs dans* $\mathscr{R}^d$ *est un processus à accroissements indépendants si quel que soit* n, $0 \leqslant t_1 < t_2 < \ldots < t_n$, *les vecteurs aléatoires* $\xi(0)$, $\xi(t_1) - \xi(0)$, ..., $\xi(t_n) - \xi(t_{n-1})$ *sont deux à deux indépendants.*

Le vecteur $\xi(0)$ est appelé *état initial* (ou *valeur initiale*) du processus, sa répartition, *répartition initiale* du processus. Un processus à accroissements indépendants au sens large est défini par la donnée de la répartition initiale $\mathsf{P}_0(B) = \mathsf{P}\{\xi(0) \in B\}$ et de la suite de répartitions $\mathsf{P}(t, h, B)$ ($t \geqslant 0$, $h > 0$, $B \in \mathfrak{B}^d$), où $\mathfrak{B}^d$ est la tribu des boréliens de $\mathscr{R}^d$, et $\mathsf{P}(t, h, B)$ la répartition du vecteur $\xi(t+h) - \xi(t)$, $\mathsf{P}(t, h, B) = \mathsf{P}\{\xi(t+h) - \xi(t) \in B\}$. En effet, si ces répartitions sont données, toute répartition conjointe des vecteurs $\xi(t_1), \xi(t_2), \ldots, \xi(t_n)$ est définie par la formule

$$\mathsf{P}\left(\bigcap_{k=0}^{n} \{\xi(t_k) \in B_k\}\right) = \int_{B_0} \mathsf{P}_0(dy_0) \int_{B_1 - y_0} \mathsf{P}(0, t_1, dy_1) \int_{B_2-(y_0+y_1)} \mathsf{P}(t_1, t_2 - t_1, dy_2) \ldots \\ \ldots \int_{B_n - (y_0 + \ldots + y_n)} \mathsf{P}(t_{n-1}, t_n - t_{n-1}, dy_n) \ldots, \quad (1)$$

où $B - z$ est l'ensemble $\{x : x = y - z,\ y \in B\}$. La répartition initiale $\mathsf{P}_0(B)$ peut être quelconque. Par ailleurs on ne peut garantir qu'à une famille de répartitions $\mathsf{P}(t, h, B)$ arbitraires est associé un processus à accroissements indépendants.

Pour que ceci ait lieu, il est nécessaire et suffisant que $\mathsf{P}(t, h, B)$ possède la propriété suivante: quels que soient n et $(t_1, \ldots, t_n)$, $t = t_0 < t_1 < \ldots < t_n = t + h$, la répartition $\mathsf{P}(t, h, B)$ est répartition de la somme des vecteurs aléatoires indépendants $\xi_1, \ldots, \xi_n$, où ξ_k admet la répartition $\mathsf{P}(t_{k-1}, t_k - t_{k-1}, B)$.

En effet, si ceci est réalisé, la famille de répartitions (1) vérifie les conditions de compatibilité.

Un processus à accroissements indépendants est *homogène* si la répartition du vecteur $\xi(t+h) - \xi(t)$ ne dépend pas de t, c'est-à-dire $\mathsf{P}(t, h, B) = \mathsf{P}(h, B)$. Les expressions et formules qui suivront sont moins compliquées pour les processus homogènes. Dans beau-

coup de cas on peut sans restreindre la généralité supposer que $\xi(0) \equiv 0$. C'est ce que nous ferons dans la suite.

Voici quelques exemples de processus à accroissements indépendants.

Processus du mouvement brownien. On appelle ainsi un processus à accroissements indépendants dont la répartition $\mathsf{P}(t, h, B)$ est gaussienne.

On sait que si l'on observe dans un puissant microscope une particule plongée dans un liquide, on constate qu'elle est animée d'un mouvement chaotique et qu'elle se déplace suivant une ligne polygonale de forme très compliquée par suite de chocs avec les molécules du liquide. Comme la particule est relativement plus grosse que les molécules, elle subit un nombre considérable de chocs en une seconde à telle enseigne qu'il est impossible d'en suivre la trajectoire. Le mouvement apparent de la particule est dit mouvement brownien. En première approximation on peut admettre que les déplacements de la particule sont indépendants et considérer le mouvement brownien comme un processus à accroissements indépendants. Comme les déplacements sont petits, on peut admettre que leur somme obéit au théorème limite central de la théorie des probabilités et que le mouvement brownien est un processus gaussien.

Quelques remarques sur la fonction de corrélation d'un processus à accroissements indépendants de moments du second ordre finis.

Soit $\xi(0) = 0$. Posons

$$a(t) = \mathsf{E}\xi(t), \quad B(t) = \mathsf{E}[\xi(t) - a(t)][\xi(t) - a(t)]^*.$$

$B(t)$ est une matrice symétrique définie non négative. Désignons $\Delta\xi(s) = \xi(t) - \xi(s)$, $\Delta a(s) = a(t) - a(s)$. Si $s < t$, on a

$$R(t, s) = \mathsf{E}[\xi(t) - a(t)][\xi(s) - a(s)]^* =$$
$$= B(s) + \mathsf{E}\Delta\xi(s)(\xi(s) - a(s))^* = B(s),$$

ou

$$R(t, s) = B(\min(t, s)). \tag{2}$$

A noter encore que

$$\mathsf{E}(\Delta\xi(s) - \Delta a(s))(\Delta\xi(s) - \Delta a(s))^* =$$
$$= B(t) - R(t, s) - R(s, t) + B(s) = B(t) - B(s). \tag{3}$$

D'où il suit en particulier que la matrice $B(t) - B(s)$ est symétrique et définie non négative.

Si les processus considérés sont homogènes, les fonctions $a(t)$ et $B(t)$ sont solutions des équations fonctionnelles suivantes:

$$a(t_1 + t_2) = a(t_1) + a(t_2), \tag{4}$$

$$B(t_1 + t_2) = B(t_1) + B(t_2). \tag{5}$$

La relation (5) est la conséquence immédiate de (3) et de l'homogénéité du processus; la relation (4) est facile à vérifier:

$$a(t_1 + t_2) = \mathsf{E}(\xi(t_1 + t_2) - \xi(t_1)) + \mathsf{E}\xi(t_2) = a(t_1) + a(t_2).$$

Si l'on admet que la fonction $a(t)$ est bornée, on sait que les solutions de l'équation fonctionnelle (4) sont de la forme $a(t) = at$, où a est un vecteur constant. La matrice $B(t)$ étant définie non négative, il s'ensuit que les solutions de l'équation (5) sont de la forme $B(t) = Bt$. Donc pour un processus homogène $\xi(t)$ à accroissements indépendants et à moments du second ordre finis ($\xi(0) = 0$) on a

$$\mathsf{E}\xi(t) = at, \quad R(t, s) = B \min(t, s). \tag{6}$$

En particulier, un mouvement brownien homogène de dimension un est défini par deux paramètres m et σ tels que $\mathsf{E}\xi(t) = mt$, $\operatorname{Var} \xi(t) = \sigma^2 t$. Si $m = 0$ et $\sigma = 1$, le mouvement brownien est un *processus wienérien*. En dimension n, un processus wienérien est un processus homogène à accroissements indépendants tel que

$$\xi(0) = 0, \quad \mathsf{E}\xi(t) = 0, \quad \mathsf{E}\xi(t)\,\xi^*(t) = It,$$

où I est la matrice unité.

Processus de Poisson. Un processus stochastiquement continu à accroissements indépendants est par définition un *processus poissonien* si, quels que soient $s, t > 0$ $(s < t)$, la répartition de la variable $\xi(t) - \xi(s)$ est poissonienne. Soit $\xi(0) = 0$. La variable $\xi(t)$ prend alors des valeurs entières et $\mathsf{P}\{\xi(t) = m\} = \frac{[a(t)]^m}{m!} e^{-at}$, $\mathsf{E}\,\xi(t) = a(t)$. On a alors

$$\mathsf{P}\{\xi(t) - \xi(s) = m\} = \frac{[a(t) - a(s)]^m}{m!} e^{-[a(t) - a(s)]}, \quad m = 0, 1, \ldots$$

Si le processus $\xi(t)$ est homogène, la monotonie de la fonction $a(t)$ entraîne

$$a(t) = at$$

et

$$\mathsf{P}\{\xi(t) - \xi(s) = m\} = \frac{a^m (t-s)^m}{m!} e^{-a(t-s)}, \quad s < t, \; m, 0, 1, \ldots$$

D'une façon générale, la continuité stochastique du processus $\xi(t)$ entraîne la continuité de la fonction $a(t)$. En effet, comme $\xi(t) \to$ $\to \xi(s)$ en probabilité lorsque $t \downarrow s$, il vient $\mathsf{E}e^{iu\xi(t)} \to \mathsf{E}e^{iu\xi(s)}$ et la fonction caractéristique $\varphi_t(u)$ de la variable $\xi(t)$ est continue en t.

D'un autre côté $\varphi_t(u) = \exp\{a(t)(e^{iu} - 1)\}$ et $\varphi_t(u)$ est continue si et seulement si la fonction $a(t)$ l'est.

Répartitions indéfiniment divisibles. Une répartition Q de $\mathcal{R}^d$ est *indéfiniment divisible* si, quel que soit n entier, Q est le n-uple produit de convolution d'une répartition $Q_n : Q = Q_n^{*n}$.

Donc, si, quel que soit n, le vecteur aléatoire ξ s'écrit $\xi = \xi_{n1} + \xi_{n2} + \ldots + \xi_{nn}$, où $\{\xi_{nk}, k = 1, \ldots, n\}$ sont des vecteurs aléatoires indépendants équirépartis, la répartition du vecteur ξ est indéfiniment divisible.

Les répartitions indéfiniment divisibles sont importantes pour les théorèmes limites de la théorie des probabilités et de la théorie des processus aléatoires. D'une part, seules les répartitions indéfiniment divisibles peuvent être répartitions limites de sommes d'infiniment petits indépendants, de l'autre, les répartitions finidimensionnelles de processus stochastiquement continus à accroissements indépendants sont indéfiniment divisibles.

Etablissons la forme générale de la fonction caractéristique $\varphi(u)$ d'une répartition indéfiniment divisible. Cette fonction s'appelle *fonction caractéristique indéfiniment divisible.*

La définition entraîne que, quel que soit n, il existe une fonction caractéristique $\varphi_n(u)$ telle que

$$\varphi(u) = [\varphi_n(u)]^n. \tag{7}$$

On conviendra que la fonction $\arg \varphi(u)$ est définie de façon univoque à l'aide des conditions: 1) $\arg \varphi(0) = 0$; 2) $\arg \varphi(u)$ est une fonction continue en u $(-\infty < u < \infty)$. Ceci posé, on peut définir de façon univoque $\ln \varphi(u)$ et $[\varphi(u)]^{1/n}$ en posant $[\varphi(u)]^{1/n} = e^{1/n \ln \varphi(u)}$. Ceci étant,

$$\lim_{n\to\infty} n(\varphi_n(u)-1) = \lim_{n\to\infty} n\left([\varphi(u)]^{\frac{1}{n}}-1\right) = \ln \varphi(u). \tag{8}$$

Théorème 1. *Soit $\varphi(h, u)$, $h \in H$, une famille de fonctions caractéristiques, H étant une suite monotone décroissante de nombres positifs tendant vers zéro. L'existence de la limite uniforme*

$$g(u) = \lim_{h\downarrow 0} \frac{\varphi(h, u)-1}{h} \tag{9}$$

sur une boule arbitraire $|u| \leqslant N$, $N > 0$, entraîne celle, sur $\{\mathcal{R}^d, \mathfrak{B}^d\}$, d'une mesure finie $\Pi(B)$, $\Pi\{0\} = 0$, d'un opérateur défini non négatif b appliquant $\mathcal{R}^d$ dans $\mathcal{R}^d$ et d'un vecteur a tels que

$$g(u) = i(a, u) - \frac{1}{2}(bu, u) + \int_{\mathcal{R}^d} \left[e^{i(u, z)} - 1 - \frac{i(u, z)}{1+|z|^2}\right] \frac{1+|z|^2}{|z|^2} \Pi(dz). \tag{10}$$

Démonstration. Soit $Q_h(\cdot)$ répartition de fonction caractéristique $\varphi(h, u)$. Posons

$$\Pi_h(B)=\frac{1}{h}\int_B \frac{|z|^2}{1+|z|^2} Q_h(dz), \quad B \in \mathfrak{B}^d.$$

On démontrera plus bas que la famille de mesures $\{\Pi_h(\cdot),\ h \in H\}$ est faiblement compacte. Choisissons une suite $h_n \downarrow 0$ telle que Π_{h_n} converge faiblement vers une mesure Π' sur $\mathfrak{B}^d$. Par ailleurs

$$\frac{\varphi(h, u)-1}{h}=\int_{\mathcal{R}^d}\left(e^{i(u, z)}-1\right)\frac{1+|z|^2}{|z|^2}\Pi_h(dz)=$$

$$=iA_h(u)-\frac{1}{2}B_h(u)+\int_{\mathcal{R}^d} f(u, z)\Pi_h(dz), \qquad (11)$$

où

$$A_h(u)=\int_{\mathcal{R}^d}\frac{(u, z)}{|z|^2}\Pi_h(dz), \qquad B_h(u)=\int_{\mathcal{R}^d}\frac{(u, z)^2}{|z|^2}\Pi_h(dz),$$

$$f(u, z)=\left(e^{i(u, z)}-1-\frac{i(u, z)}{1+|z|^2}+\frac{1}{2}\frac{(u, z)^2}{1+|z|^2}\right)\frac{1+|z|^2}{|z|^2}.$$

Si l'on admet que $f(u, 0)=0$, alors $f(u, z)$ sera une fonction continue et bornée. Donc

$$\lim \int_{\mathcal{R}^d} f(u, z)\Pi_{h_n}(dz)=\int_{\mathcal{R}^d} f(u, z)\Pi'(dz).$$

L'existence de la limite du premier membre de (11), lorsque $h=h_n$ et $n\to\infty$, entraîne celle des limites

$$\lim A_{h_n}(u)=a(u), \qquad \lim B_{h_n}(u)=B(u),$$

où $a(u)$ est une fonction linéaire, $B(u)$ une forme quadratique définie positive, c'est-à-dire $a(u)=(a, u)$ et $B(u)=(b'u, u)$, où b' est un opérateur symétrique défini non négatif. En passant à la limite sur h_n dans (11), on obtient

$$g(u)=i(a, u)-\frac{1}{2}(b'u, u)+\int_{\mathcal{R}^d} f(u, z)\Pi'(dz). \qquad (12)$$

Soit $\Pi(A)=\Pi'(A-\{0\})$ ($\{0\}$ est un ensemble constitué du seul point 0). Dans l'intégrale de (12) on peut remplacer la mesure $\Pi'(\cdot)$ par la mesure $\Pi(\cdot)$. Par ailleurs l'intégrale

$$\frac{1}{2}\int_{\mathcal{R}^d}\frac{(u, z)^2}{|z|^2}\Pi(dz)$$

existe et constitue une forme quadratique définie non négative $(b''u, u)$. On remarque aisément que $(b'u, u)\geqslant(b''u, u)$. Donc $b=$

$= b' - b''$ est un opérateur symétrique défini non négatif. Par suite

$$g(u) = i(a, u) - \frac{1}{2}(bu, u) + \int_{\mathscr{R}^d} \left(f(u, z) - \frac{1}{2} \frac{(u, z)^2}{|z|^2} \right) \Pi(dz). \quad \blacksquare$$

Passons maintenant à la démonstration de la faible compacité de la famille $\{\Pi_h, h \in H\}$. Il nous faut prouver que

$$\text{a) } \Pi_h(\mathscr{R}^d) \leqslant L; \quad \text{b) } \lim_{N\to\infty} \overline{\lim_{h\downarrow 0}} \Pi_h(\overline{K}_N) = 0;$$

K_N représente le cube fermé

$$K_N = \{z = (z^1, \ldots, z^d) : \max_j |z^j| \leqslant N\}$$

et $\overline{K}_N = \mathscr{R}^d \setminus K_N$.

Supposons que $|u| \leqslant N_1$, N_1 étant arbitraire. Par hypothèse et de (11) il résulte que

$\forall \delta > 0$, $\exists h_0 = h_0(N_1, \delta)$ tel que pour tout $c > 0$

$$-\operatorname{Re} g(u) + \delta \geqslant \int_{K_c} \frac{1 - \cos(u, z)}{|z|^2} \Pi_h(dz), \quad h \leqslant h_0,$$

et pour $c > 1$

$$-\operatorname{Re} g(u) + \delta \geqslant \int_{\overline{K}_c} (1 - \cos(u, z)) \Pi_h(dz).$$

Intégrons ces inégalités sur $u \in K_\rho$ et divisons par le volume de K_ρ. On obtient

$$-(2\rho)^{-d} \int_{K_\rho} \operatorname{Re} g(u)\, du + \delta \geqslant \int_{K_c} \frac{1}{|z|^2} \left(1 - \prod_{k=1}^{m} \frac{\sin \rho z^k}{\rho z^k}\right) \Pi_h(dz) \quad (13)$$

et

$$-(2\rho)^{-d} \int_{K_\rho} \operatorname{Re} g(u)\, du + \delta \geqslant \int_{\overline{K}_c} \left(1 - \prod_{k=1}^{m} \frac{\sin \rho z^k}{\rho z^k}\right) \Pi_h(dz). \quad (14)$$

Comme $\frac{\sin x}{x} \geqslant 1 - \frac{x^2}{6}$ pour tous les $x \geqslant 0$ et $1 - \prod_{k=1}^{m} (1 - \alpha_k) \geqslant$ $\geqslant \sum_{k=1}^{m} \alpha_k - \sum_{k \neq j} \alpha_k \alpha_j$ pour tous les $\alpha_k \in [0, 1]$, il vient

$$\frac{1}{|z|^2} \left(1 - \prod_{k=1}^{m} \frac{\sin \rho z^k}{\rho z^k}\right) \geqslant \frac{\rho^2}{6} \left(1 - \frac{\rho^2 |z|^2}{6}\right) \geqslant \frac{\rho^2}{6} \left(1 - \frac{\rho^2 d c^2}{6}\right).$$

En faisant $\rho^2 = \frac{3}{dc^2}$, on déduit de (13)

$$\Pi_h(K_c) \leqslant 4dc^2 \Big[\delta - (2\rho)^{-d} \int_{K_c} \operatorname{Re} g(u)\,du\Big] \quad \forall h \in]0, h_0]. \qquad (15)$$

En remarquant par ailleurs que $\left|\prod_{k=1}^{d} \frac{\sin \rho z^k}{\rho z^k}\right| \leqslant \frac{1}{\rho c}$ pour $z \overline{\in} K_c$, on déduit de (14) ($\rho c > 1$)

$$\Pi_h(\overline{K}_c) \leqslant \frac{\rho c}{\rho c - 1}\Big[\delta - (2\rho)^{-d} \int_{K_\rho} \operatorname{Re} g(u)\,du\Big] \quad \forall h \in]0, h_0]. \qquad (16)$$

Donc $\Pi_h(\mathscr{R}^d) = \Pi_h(K_c) + \Pi_h(\overline{K}_c) \leqslant L$, où la constante L ne dépend pas de $h \in]0, h_0]$. On remarquera maintenant que par hypothèse la fonction $g(u)$ est continue et que $g(0) = 0$. Donc $\forall \delta > 0 \quad \exists \rho_1$ assez petit tel que

$$\left|(2\rho_1)^{-d} \int_{K_{\rho_1}} \operatorname{Re} g(u)\,du\right| < \delta.$$

L'inégalité (16) entraîne pour c suffisamment grand

$$\Pi_h(\overline{K}_c) < 2\delta$$

$\forall h \in]0, h_0]$. Ceci démontre la compacité de la famille de mesures $\{\Pi_h, h \in H\}$. ■

Ce théorème et la formule (10) entraînent immédiatement le

THÉORÈME 2. *La fonction caractéristique $\varphi(u)$ d'une répartition indéfiniment divisible dans $\mathscr{R}^d$ est de la forme*

$$\varphi(u) = \exp\{g(u)\},$$

où $g(u)$ est définie par (10).

Mettons la formule (10) sous une autre forme. Supposons que $c > 0$. Les intégrales

$$\int_{S_c} (u, z)\,\Pi(dz), \qquad \int_{\overline{S}_c} \frac{(u, z)}{|z|^2}\,\Pi(dz)$$

étant finies (S_c est la sphère de rayon $c > 0$ centrée en l'origine des coordonnées), on a

$$\int_{\mathscr{R}^d} \Big(e^{i(u,z)} - 1 - \frac{i(u, z)}{1 + |z|^2}\Big)\Pi(dz) =$$

$$= \int_{S_c} (e^{i(u,z)} - 1 - i(u, z))\,\overline{\Pi}(dz) + \int_{\overline{S}_c} (e^{i(u,z)} - 1)\,\overline{\Pi}(dz) + i(u, a'),$$

où

$$a' = \int_{S_c} z\Pi(dz) - \int_{\bar{S}_c} \frac{z}{|z|^2}\Pi(dz), \quad \bar{\Pi}(A) = \int_A \frac{1+|z|^2}{|z|^2}\Pi(dz).$$

Si l'on pose $a_c = a + a'$, on obtient la représentation suivante pour $g(u)$:

$$g(u) = i(a_c, u) - \frac{1}{2}(bu, u) + \\ + \int_{S_c} (e^{i(u, z)} - 1 - i(u, z))\bar{\Pi}(dz) + \int_{\bar{S}_c} (e^{i(u, z)} - 1)\bar{\Pi}(dz). \quad (17)$$

Contrairement à la mesure $\Pi(A)$, la mesure $\bar{\Pi}(A)$ n'est plus d'une façon générale finie, mais $\bar{\Pi}(\bar{S}_c) < \infty$ quel que soit $c > 0$ et $\bar{\Pi}\{0\} = 0$. Par ailleurs on démontrera prochainement que la mesure $\bar{\Pi}(A)$ admet une interprétation probabiliste plus simple que la mesure $\Pi(A)$.

Supposons que les répartitions $Q_h(\cdot)$ de fonctions caractéristiques $\varphi(h, u)$ admettent un moment du second ordre fini. Substituons aux mesures $\Pi_h(B)$ les mesures finies

$$G_h(B) = \frac{1}{h}\int_B |z|^2 Q_h(dz).$$

Par hypothèse $\forall N_1 > 0$ et $\forall \delta > 0$, $\exists h_0 = h_0(N_1, \delta)$ tel que pour tout $h \in]0, h_0[$

$$-\operatorname{Re} g(u) + \delta \geqslant \int_{\mathcal{R}^d} \frac{1 - \cos(u, z)}{|z|^2} G_h(dz)$$

pour tous les u, $|u| \leqslant N_1$. On peut donc reprendre tous les raisonnements qui ont servi à démontrer le théorème 1 et établir la faible compacité de la famille de mesures $G_h(\cdot)$. La relation

$$\frac{\varphi(h, u) - 1}{h} = i\tilde{A}_h(u) - \frac{1}{2}\tilde{B}_h(u) + \int_{\mathcal{R}^d} \tilde{f}(u, z)G_h(dz),$$

où

$$\tilde{f}(u, z) = \left(e^{i(u, z)} - 1 - i(u, z) + \frac{1}{2}(u, z)^2\right)\frac{1}{|z|^2},$$

$$\tilde{A}_h u = \int_{\mathcal{R}^d} \frac{(u, z)}{|z|^2} G_h(dz), \quad \tilde{B}_h(u) = \int_{\mathcal{R}^d} \frac{(u, z)^2}{|z|^2} G_h(\ z),$$

entraîne comme dans la démonstration du théorème 1

$$g(u)=i(\widetilde{a}, u)-\frac{1}{2}(\widetilde{b}u, u)+\int_{\mathcal{R}^d}[e^{i(u,z)}-1-i(u, z)]\frac{1}{|z|^2}G(dz), \tag{18}$$

où $G(\cdot)$ est une mesure finie sur $\mathcal{R}^d$ et $G\{0\}=0$.

Nous avons ainsi obtenu le complément suivant au théorème 1.

THÉORÈME 3. *Si dans les hypothèses du théorème* 1 *les répartitions* $Q_h(\cdot)$ *possèdent en outre des moments finis du second ordre, alors la fonction* $g(u)$ *admet la représentation* (18).

Appliquons les théorèmes 1 et 3 aux processus stochastiquement continus homogènes à accroissements indépendants.

THÉORÈME 4. *Soit* $\varphi(t, u)$ *fonction caractéristique du vecteur* $\xi(s+t)-\xi(s)$, $s \geqslant 0$, $t>0$, *où* $\xi(t)$ *est un processus stochastiquement continu homogène à accroissements indépendants à valeurs dans* $\mathcal{R}^d$. *Alors*

$$\varphi(t, u)=\exp\{tg(u)\}, \tag{19}$$

où $g(u)$ *est défini par* (10) *(ou* (17)). *Si le processus* $\xi(t)$ *possède des moments finis du second ordre, la fonction* $g(u)$ *admet la représentation* (18).

D é m o n s t r a t i o n. Comme

$$|\varphi(t, u)-\varphi(s, u)| \leqslant \mathsf{E}|e^{i(u, \xi(t)-\xi(s))}-1|,$$

la continuité stochastique du processus $\xi(t)$ entraîne la continuité de la fonction $\varphi(t, u)$ en t. D'un autre côté, l'homogénéité du processus $\xi(t)$ et l'indépendance de ses accroissements donnent

$$\begin{aligned}\varphi(t_1+t_2, u) &= \mathsf{E}\exp\{i(u, \xi(t_1+t_2)-\xi(t_1))+i(u, \xi(t_1)-\xi(0))\}= \\ &= \mathsf{E}\exp\{i(u, \xi(t_2)-\xi(0))\}\,\mathsf{E}\exp\{i(u, \xi(t_1)-\xi(0))\}= \\ &= \varphi(t_1, u)\cdot\varphi(t_2, u).\end{aligned}$$

Or l'équation $f(t+s)=f(t)f(s)$ $(t \geqslant 0,\ s \geqslant 0)$ admet une solution continue unique de la forme $f(t)=e^{at}$. Donc $\varphi(t, u)=e^{tg(u)}$, où $g(u)=\lim\limits_{t \downarrow 0}\frac{\varphi(t, u)-1}{t}$. D'où et en vertu des théorèmes 1 et 3 suit le théorème annoncé. ■

Voyons quelques cas particuliers de la formule (19) dans laquelle $g(u)$ sera définie par (17).

a) $b=0$, $\overline{\Pi}(\cdot) \equiv 0$.

Dans ce cas $\varphi(t, u)=e^{it(a, u)}$: c'est la fonction caractéristique d'une répartition dégénérée concentrée au point ta. Donc $\xi(t)=$

$= \xi(0) + at$ et le point $\xi(t)$ est animé d'un mouvement uniforme de vitesse a.

b) $\overline{\Pi}(\cdot) \equiv 0$.

L'accroissement $\xi(t+s) - \xi(s)$ possède alors une répartition normale de moyenne at et de matrice de corrélation bt. Si $\xi(0) = 0$, le processus $\xi(t)$ est gaussien, c'est-à-dire est un processus de mouvement brownien.

c) $a = 0$, $b = 0$, la mesure $\overline{\Pi}$ est concentrée au point z_0, $|z_0| < c$, et $\overline{\Pi}\{z_0\} = q$.

On a

$$\varphi(t, u) = \exp\{qt(e^{i(u, z_0)} - 1 - i(u, z_0))\}.$$

On vérifie aisément que l'accroissement $\xi(t)$ ($\xi(0) = 0$) peut se mettre sous la forme

$$\xi(t) = z_0(\nu(t) - qt),$$

où $\nu(t)$ est un processus poissonien homogène de paramètre q, $\mathsf{E}\nu(t) = qt$.

d) $b = 0$, $\overline{\Pi}(S_c) < \infty$. La formule (17) s'écrit alors

$$g(u) = i(\bar{a}, u) + q \int_{\mathscr{R}^d} (e^{i(u, z)} - 1)\Pi_0(du),$$

où Π_0 est une mesure probabiliste sur $\mathfrak{B}^d$ ($q = \overline{\Pi}(\mathscr{R}^d)$). L'expression

$$\varphi(t, u) = e^{i(\bar{a}t, u)} \sum_{n=0}^{\infty} e^{-qt} \frac{(qt)^n}{n!} \left[\int_{\mathscr{R}^d} e^{i(u, z)} \Pi_0(dz) \right]^n$$

admet une interprétation simple. C'est la fonction caractéristique de la somme $\bar{a}t + \xi_1 + \ldots + \xi_{\nu(t)}$, où $\xi_1, \ldots, \xi_n, \ldots$ sont des vecteurs indépendants dans $\mathscr{R}^d$ de même répartition $\Pi_0(\cdot)$, $\nu(t)$ étant un processus poissonien de paramètre q, ne dépendant pas de $\xi_1, \ldots, \xi_n, \ldots$

Le processus $\xi(t)$ s'appelle *processus poissonien général*.

e) En appliquant la formule de Khintchine, on peut mettre la fonction caractéristique d'un processus homogène unidimensionnel à accroissements indépendants sous la forme

$$\mathsf{E}e^{iu\xi(t)} = \exp\left\{t\left(i\gamma u + \int_{-\infty}^{\infty}\left(e^{iux} - 1 - \frac{iux}{1+x^2}\right)\frac{1+x^2}{x^2}\,dG(x)\right)\right\},$$

où γ est un nombre réel, G une fonction non croissante bornée sur $]-\infty, \infty[$, l'expression $\left(e^{iux} - 1 - \frac{iux}{1+x^2}\right)\frac{1+x^2}{x^2}$ est définie par continuité pour $x = 0$ et vaut $-\frac{1}{2}u^2$.

Pour expliciter la fonction caractéristique d'un processus stochastiquement continu arbitraire à accroissements indépendants on se servira du théorème suivant de Khintchine [2].

THÉORÈME 5. *Soient* $\xi_{n1}, \ldots, \xi_{nm_n}$, $n = 1, 2, \ldots$, *des vecteurs aléatoires indépendants (pour n donné) tels que*

$$\lim_{n\to\infty} \max_{1\leqslant k\leqslant m_n} \mathsf{P}\{|\xi_{nk}|>\varepsilon\}=0 \quad \forall\varepsilon>0.$$

Si les répartitions de la suite de vecteurs aléatoires $\zeta_n = \xi_{n1} + \xi_{n2} + \ldots \ldots + \xi_{nm_n}$ *sont faiblement convergentes, alors la répartition limite est indéfiniment divisible.*

Dans le cas unidimensionnel on trouve la démonstration par exemple dans B. Gnédenko et A. Kolmogorov [1] et dans V. Pétrov [1]. La généralisation ne présente aucune difficulté.

Du théorème 5 on déduit la formule suivante pour la fonction caractéristique $\varphi(t, u)$ du processus stochastiquement continu à accroissements indépendants $\xi(t)$ ($\xi(0) = 0$):

$$\varphi(t, u)=\exp\Big\{i(a(t), u)-\frac{1}{2}(b(t)u, u)+ + \int_{\mathcal{R}^d}\Big[e^{i(u, z)}-1-\frac{i(u, z)}{1+|z|^2}\Big]\frac{1+|z|^2}{|z|^2}\Pi(t, dz)\Big\}, \tag{20}$$

la fonction caractéristique de l'accroissement $\xi(t) - \xi(s)$, $0 \leqslant s < < t$, peut s'écrire

$$\varphi(s, t; u) = \exp\{g(t, u) - g(s, u)\}. \tag{21}$$

§ 4. Processus markovien au sens large

La notion de processus markovien repose sur le principe de processus sans postaction. Imaginons-nous un système (une particule) susceptible de se trouver dans des états différents. Les états possibles du système forment un ensemble X appelé *espace des phases.* Convenons que le système évolue dans le temps. Soit x_t son état à l'instant t. Si $x_t \in B$, où $B \subset X$, on dit que le système se trouve dans l'ensemble B à l'instant t. Supposons que l'évolution du système est stochastique, c'est-à-dire l'état du système à l'instant t d'une façon générale n'est pas défini univoquement en fonction des états pris par le système à des instants s antérieurs ($s < t$), mais est aléatoire et obéit à certaines lois probabilistes. Soit $\mathsf{P}(s, x, t, B)$ la probabilité de l'événement $x_t \in B$ ($s < t$) sachant que $x_s = x$.

La fonction $\mathsf{P}(s, x, t, B)$ est appelée *probabilité de passage* ou *de transition* du système envisagé. Par système sans postaction on entend un système pour lequel la probabilité de se trouver à l'instant t dans l'ensemble B, le mouvement du système étant entièrement

connu jusqu'à l'instant s ($s < t$), est encore égale à $\mathsf{P}(s, x, t, B)$ et par conséquent ne dépend que de l'état du système au dernier instant connu. Cette définition sera entièrement formalisée au chapitre VII. Pour l'instant nous allons donner de cette notion une définition simple et suffisante pour de nombreux problèmes.

Soit $\mathsf{P}(s, x, u, y, t, B)$ la probabilité conditionnelle de l'événement $x_t \in B$ sachant que $x_s = x$, $x_u = y$ ($s < u < t$). En vertu des propriétés générales des probabilités conditionnelles on a

$$\mathsf{P}(s, x, t, B) = \int_X \mathsf{P}(s, x, u, y, t, B)\,\mathsf{P}(s, x, u, dy). \tag{1}$$

S'agissant d'un système sans postaction, il est naturel de supposer que $\mathsf{P}(s, x, u, y, t, B) = \mathsf{P}(u, y, t, B)$. Dans ce cas l'égalité (1) s'écrit

$$\mathsf{P}(s, x, t, B) = \int_X \mathsf{P}(u, y, t, B)\,\mathsf{P}(s, x, u, dy). \tag{2}$$

La relation (2) s'appelle *équation de Kolmogorov-Chapman*. Elle peut servir de base à la définition du processus sans postaction connu sous le nom de processus markovien.

Soit $\{X, \mathfrak{B}\}$ un ensemble mesurable. On appellera *noyau stochastique* une fonction $\mathsf{P}(x, B)$, $x \in X$, $B \in \mathfrak{B}$, vérifiant les conditions suivantes:

a) $\mathsf{P}(x, B)$ est une mesure sur $\mathfrak{B}$ pour x fixe et $\mathsf{P}(x, X) = 1$,

b) $\mathsf{P}(x, B)$ est une fonction $\mathfrak{B}$-mesurable de x pour B fixe.

Nous adopterons cette terminologie pour le cas plus général où l'argument x de la fonction $\mathsf{P}(x, B)$ prend ses valeurs dans un ensemble mesurable $\{X_0, \mathfrak{B}_0\}$ distinct de $\{X, \mathfrak{B}\}$.

Soit I un intervalle semi-ouvert (fermé) fini ou non. On appelle *famille markovienne de noyaux stochastiques* une famille de noyaux stochastiques $\{\mathsf{P}_{st}(x, B) = \mathsf{P}(s, x, t, B),\ s < t,\ (s, t) \in I \times I\}$ vérifiant l'équation de Kolmogorov-Chapman.

DÉFINITION. *Un processus markovien au sens large est par définition un triplet composé*

a) *d'un espace mesurable* $\{X, \mathfrak{B}\}$,

b) *d'un intervalle semi-ouvert (fermé) I de l'axe réel*,

c) *d'une famille markovienne de noyaux stochastiques*

$$\{\mathsf{P}_{st}(x, B),\ s < t,\ (s, t) \in I \times I\}.$$

La famille de noyaux $\mathsf{P}_{st}(x, B) = \mathsf{P}(s, x, t, B)$ est appelée *probabilité de passage* du processus markovien, l'espace $\{X, \mathfrak{B}\}$ *espace des phases* du système, les points de l'ensemble I représentent le temps, $\mathsf{P}_{st}(x, B) = \mathsf{P}(s, x, t, B)$ la probabilité conditionnelle que le système se trouve dans l'ensemble B à l'instant t sachant qu'à l'instant s il se trouvait en un point x de l'espace des phases ($s < t$).

Dans la suite on supposera que le noyau $\mathsf{P}_{st}(x, B)$ est défini aussi pour $s = t$. Plus exactement on posera

$$\mathsf{P}_{tt}(x, B) = \chi(B, x),$$

où $\chi(B, x)$ est l'indicateur de l'ensemble B, c'est-à-dire

$$\chi(B, x) = \begin{cases} 1 \text{ si } x \in B, \\ 0 \text{ si } x \bar{\in} B. \end{cases}$$

Il est évident qu'avec cette définition du noyau $\mathsf{P}_{tt}(x, B)$, l'équation (2) reste toujours valable pour $u = s$ ou $u = t$.

Opérateurs engendrés par des probabilités de passage. Les probabilités de passage peuvent être associées à deux familles d'opérateurs.

Soit $\mathscr{M} = \mathscr{M}(\mathfrak{B})$ l'ensemble de toutes les mesures finies sur $\mathfrak{B}$. Posons $m_{ts} = T_{ts}^* m$, où

$$m_{ts}(B) = \int \mathsf{P}(s, y, t, B)\, m(dy), \quad s \leqslant t, \quad B \in \mathfrak{B}. \tag{3}$$

Si la mesure $m(\cdot)$ est la répartition du système à l'instant s, $m(B) = \mathsf{P}\{x_s \in B\}$, la formule (3) est la « formule de la probabilité totale » et $m_{ts}(\cdot)$ la répartition du système à l'instant t, $m_{ts}(B) = \mathsf{P}\{x_t \in B\}$. Donc l'opérateur T_{ts}^* exprime la répartition du système envisagé dans son espace des phases à l'instant t en fonction de la répartition à l'instant s, $s < t$.

Il est évident que T_{ts}^* est un opérateur de $\mathscr{M}$ dans $\mathscr{M}$. De la formule de Kolmogorov-Chapman on déduit une loi simple de composition des opérateurs T_{ts}^*.

Soit $s < u < t$. En se servant de l'équation (2) et en modifiant l'ordre d'intégration dans les intégrales multiples, on obtient

$$m_{ts}(B) = \int_X m(dx) \int_X \mathsf{P}(u, y, t, B)\, \mathsf{P}(s, x, u, dy) =$$

$$= \int_X \left(\int_X \mathsf{P}(s, x, u, dy)\, m(dx) \right) \mathsf{P}(u, y, t, B) =$$

$$= \int_X m_{us}(dy)\, \mathsf{P}(u, y, t, B).$$

Par suite

$$T_{ts}^* = T_{tu}^* T_{u}^* \quad (s < u < t). \tag{4}$$

Donc la famille d'opérateurs T_{ts}^* envisagée comme fonction de l'intervalle $[s, t]$ est une fonction d'intervalle multiplicative orientée. Par orientation d'une fonction d'intervalle multiplicative on entendra l'un des deux modes de disposition des facteurs correspondant à une partition de l'intervalle.

La deuxième famille d'opérateurs considérée agit dans l'espace de toutes les fonctions $\mathfrak{B}$-mesurables bornées. Soient $\mathscr{B}(\mathfrak{B})$ cette famille et $\mathscr{B}(\mathfrak{B})_+$ l'ensemble de toutes les fonctions non négatives de cet espace. Posons $f_{st} = T_{st}f$ si

$$f_{st}(x) = \int f(y)\, \mathsf{P}(s, x, t, dy).$$

La fonction $f_{st}(x)$ admet l'interprétation probabiliste évidente suivante. Elle est égale à l'espérance mathématique de la variable aléatoire $f(x_t)$ sachant qu'à l'instant s $(s < t)$ le système se trouvait dans l'état x $(x_s = x)$. La $\mathfrak{B}$-mesurabilité de la probabilité de passage en x entraîne celle de la fonction $f_{st}(x)$. Munissons $\mathscr{B}(\mathfrak{B})$ de la norme $\|f\| = \sup |f(x)|$. De toute évidence $\|f_{st}\| \leqslant \|f\|$.

Donc les opérateurs T_{st} appliquent $\mathscr{B}(\mathfrak{B})$ et $\mathscr{B}(\mathfrak{B})_+$ dans eux-mêmes. En se servant de la formule de Kolmogorov-Chapman et en modifiant l'ordre d'intégration, on obtient pour $s < u < t$

$$f_{st}(x) = \int f(y)\, \mathsf{P}(s, x, t, y) = \int f(y) \int \mathsf{P}(s, x, u, dz)\, \mathsf{P}(u, z, t, dy) =$$
$$= \int f_{ut}(z)\, \mathsf{P}(s, x, u, dz),$$

ou

$$T_{st} = T_{su}T_{ut} \quad (s < u < t). \tag{5}$$

Les opérateurs T_{st} forment donc une fonction d'intervalle multiplicative (non commutative) qui n'est pas orientée de la même manière que T^*_{st}. De toute évidence, $T_{st}1 = 1$ et $T_{ss}f = f$ et $T_{ss} = I$, où I est l'opérateur unité.

Définition. *Un processus markovien est homogène si $I = [0, \infty[$ et les noyaux $\mathsf{P}_{st}(x, B)$ dépendent comme fonctions des arguments (s, t) uniquement de la différence $t - s$:*

$$\mathsf{P}_{st}(x, B) = \mathsf{P}_{t-s}(x, B) \quad (t > s).$$

Dans le cas de processus markoviens homogènes l'équation de Kolmogorov-Chapman s'écrit

$$\mathsf{P}_{u+v}(x, B) = \int \mathsf{P}_u(x, dy)\, \mathsf{P}_v(y, B), \quad u > 0, \ v > 0. \tag{6}$$

La famille de noyaux $\{\mathsf{P}_t(x, B), t \geqslant 0\}$ s'appelle également probabilité de passage d'un processus markovien homogène.

Dans le cas homogène les opérateurs $T^*_{t+s,\ s}$, $T_{s,\ s+t}$ ne dépendent pas de s et au lieu des familles d'opérateurs à deux paramètres $\{T^*_{ts}, t > s > 0\}$, $\{T_{st}, 0 < s < t\}$ il semble plus logique de considérer les familles à un paramètre $\{T^*_t, t \geqslant 0\}$, $\{T_t, t > 0\}$ défi-

nies par

$$T_t^* m(B) = \int \mathsf{P}_t(x, B)\, m(dx),$$

$$T_t f(x) = \int f(y)\, \mathsf{P}_t(x, dy).$$

Les produits de convolution (4) et (5) s'écrivent

$$T_{u+v}^* = T_u^* T_v^*, \quad T_{u+v} = T_u T_v.$$

Ceci traduit le fait que les familles d'opérateurs $\{T_t^*, t > 0\}$, $\{T_t, t > 0\}$ forment un sous-groupe d'opérateurs dans les espaces respectifs.

Equations de Kolmogorov. Parmi les problèmes majeurs de la théorie des processus markoviens au sens large on distinguera :

a) le choix des plus importantes classes (modèles) de processus markoviens possédant des propriétés spécifiques et leur description en termes de probabilités de passage ;

b) la description constructive des probabilités de passage correspondant à la classe donnée de processus markoviens ;

c) l'étude des propriétés asymptotiques (limites) des probabilités de passage des différentes classes de processus markoviens.

Il va de soi que ces formulations sont très générales et très vagues, et le programme défini très vaste. Dans le présent paragraphe on n'exhibera que des résultats préliminaires.

Le premier critère de classification des processus markoviens est l'espace des phases. Dans cette optique, les processus markoviens les plus simples sont les processus à nombre d'états fini ou dénombrable. Dans le dernier cas, en assujettissant la probabilité de passage à certaines contraintes analytiques on peut linéariser l'équation de Kolmogorov-Chapman pour obtenir des systèmes d'équations différentielles ordinaires (dites équations différentielles de Kolmogorov directes et inverses) qui déterminent entièrement la probabilité de passage dans de nombreux cas.

Si l'on se place dans des espaces des phases plus généraux, on peut définir des classes de processus markoviens au sens large en attribuant à leurs probabilités de passage des propriétés traduisant des idées intuitives sur le caractère du mouvement du système dans l'espace des phases. Dans cet optique on envisage les classes suivantes de processus :

a) Processus discontinus. Un processus est discontinu si le système change d'état à n'importe quel instant par saut d'amplitude aléatoire.

Dans ces processus on distinguera ceux dont l'état reste constant entre les sauts et ceux qui varient de façon déterminée entre les sauts.

b) Processus de diffusion. On appelle ainsi les processus définis sur des espaces vectoriels finidimensionnels se comportant sur de petits intervalles de temps comme des processus du mouvement brownien.

c) Processus markoviens sur un espace finidimensionnel, approchés sur de petits intervalles de temps par un processus arbitraire à accroissements indépendants.

Pour définir les diverses classes de processus markoviens au sens large on s'appuie généralement sur l'idée de la linéarisation de l'équation de Kolmogorov-Chapman, c'est-à-dire on impose à la probabilité de passage des conditions permettant de passer des équations non linéaires (2) à des équations linéaires. Ces équations peuvent être des équations intégro-différentielles ou bien des équations aux dérivées partielles de type parabolique ou bien encore des équations contenant aussi bien des dérivées partielles du premier et du second ordre que des intégrales. Voyons quelques généralités sur la déduction de ces équations.

Soit $I = [0, t^*[$. Fixons un $t \in I$ et considérons la classe $\mathcal{D}$ de fonctions de $\mathscr{B}(\mathfrak{B})$ telles que pour $s \in]0, t[$ et $x \in X$ existent les limites

$$\lim_{h \downarrow 0} \frac{T_{s-h,\, s} f(x) - f(x)}{h} = A_s f(x)$$

et

$$\lim_{h \downarrow 0} T_{t-h,\, t} f(x) = f(x).$$

A_s est un opérateur défini sur $\mathcal{D}$ dépendant de $s \in]0, t[$ considéré comme paramètre. De toute évidence $\mathcal{D}$ est un espace vectoriel et A_s un opérateur linéaire.

Posons

$$f(s, x) = T_{st} f(x), \quad s \in [0, t].$$

Supposons que $T_{st} f(x) \in \mathcal{D}$. La dérivée à gauche par rapport à s de $f(s, x)$ s'écrit

$$\frac{\partial f(s, x)}{\partial s} = \lim_{h \downarrow 0} \frac{f(s-h, x) - f(s, x)}{-h} =$$

$$= -\lim_{h \downarrow 0} \frac{T_{s-h,s} f(s, x) - f(s, x)}{h} = -A_s f(s, x).$$

Cette relation entraîne dans beaucoup de cas l'existence de la dérivée $\frac{\partial f(s,x)}{\partial s}$ qui est donc solution de l'équation

$$\frac{\partial f(s, x)}{\partial s} = -A_s f(s, x), \quad s \in]0, t[. \tag{7}$$

Par définition de l'ensemble $\mathscr{D}$ il faut ajouter à cette équation la condition aux limites

$$\lim_{s \uparrow t} f(s, x) = f(x). \tag{8}$$

Si $\chi(B, x) \in \mathscr{D}$, en posant $f(x) = \chi(B, x)$ on remarque que $f(s, x) = \mathsf{P}(s, x, t, B)$ est solution des équations

$$\frac{\partial \mathsf{P}(s, x, t, B)}{\partial s} = -A_s [\mathsf{P}(s, x, t, B)], \quad s \in [0, t],$$

$$\lim_{s \uparrow t} \mathsf{P}(s, x, t, B) = \chi(B, x).$$

Même si $\chi(B, x) \bar{\in} \mathscr{D}$, la classe de fonctions $\mathscr{D}$ peut s'avérer suffisamment riche pour que les valeurs $T_{st} f(x)$ $(f \in \mathscr{D})$ définissent univoquement les probabilités de passage $\mathsf{P}(s, x, t, B)$. Ceci a lieu lorsque la classe $\mathscr{D}$ est partout dense dans $\mathscr{B}(\mathfrak{B})$ pour la convergence ponctuelle bornée des fonctions. Ceci étant, si le théorème d'existence et d'unicité de la solution des équations (7) et (8) est valable pour toute fonction $f \in \mathscr{D}$, ces équations définissent de façon univoque les probabilités de passage (sur l'intervalle $]0, t[$) et peuvent être utilisées pour la détermination de la fonction $\mathsf{P}(s, x, t, B)$ ou pour l'étude de ses propriétés.

L'équation (7) est appelée *première équation* (ou *équation inverse*) *de Kolmogorov.*

On peut généraliser ces raisonnements à la famille d'opérateurs $\{T_{ts}^*, t \geqslant s \geqslant 0\}$.

Soient $\mathscr{W} = \mathscr{W}(\mathfrak{B})$ l'espace de toutes les charges finies sur $\mathfrak{B}$, c'est-à-dire l'ensemble de toutes les fonctions d'ensembles dénombrablement additives finies sur $\mathfrak{B}$, et $\mathscr{D}^*$ le sous-ensemble de $\mathscr{W}$ composé de toutes les charges $q(B)$ telles que

$$\lim_{h \downarrow 0} \frac{T_{t+h,\, t} q(B) - q(B)}{h} = A_t^* q(B),$$

$$\lim_{h \downarrow 0} T_{s+h,\, s} q(B) = q(B)$$

quels que soient $(t, B) \in [s, t^*[\times \mathfrak{B}_0$, où $\mathfrak{B}_0 \subset \mathfrak{B}$, s étant fixe.

Posons $q(t, B) = T_{ts}^* q(B)$. Si la charge $q(B)$ est telle que $q(t, B) \in \mathscr{D}^*$, alors la dérivée première $\frac{\partial q(t, B)}{\partial t}$ existe et

$$\frac{\partial q(t, B)}{\partial t} = \lim_{h \downarrow 0} \frac{q(t+h, B) - q(t, B)}{h} = \lim_{h \downarrow 0} \frac{T_{t+h,\, t} q(t, B) - q(t, B)}{h},$$

de sorte que $q(t, B)$ est solution des équations

$$\frac{\partial q(t, B)}{\partial t} = A_t^* q(t, B), \tag{9}$$

$$\lim_{t \downarrow s} q(t, B) = q(B). \tag{10}$$

Si $\chi(B, x) \in \mathcal{I}^*$, les équations (9), (10) deviennent pour $q(B) = \chi(B, x)$

$$\frac{\partial \mathsf{P}(s, x, t, B)}{\partial t} = A_t^*[\mathsf{P}(s, x, t, B)], \quad \lim_{t \downarrow s} \mathsf{P}(s, x, t, B) = \chi(B, x).$$

L'équation (9) porte le nom de *deuxième équation* (ou *équation directe*) *de Kolmogorov.* On peut tenir les mêmes raisonnements que pour l'équation inverse.

Dans la suite on définira la forme des opérateurs A_s et A_t^* pour des classes particulières de processus markoviens.

Processus à nombre fini ou dénombrable d'états. Soient X un espace constitué d'un nombre fini ou dénombrable de points, $\mathfrak{B}$ l'ensemble des parties de X. On désignera par $i, j, k, \ldots$ les points de X. Soit un processus markovien au sens large à valeurs dans X. Posons

$$p_{ij}(s, t) = \mathsf{P}(s, i, t, \{j\}).$$

Les probabilités $p_{ij}(s, t)$ définissent de toute évidence la probabilité de passage pour un ensemble quelconque $B \subset X$

$$\mathsf{P}(s, i, t, B) = \sum_{j \in B} p_{ij}(s, t).$$

Les relations suivantes sont immédiates:

$$p_{ij}(s, t) \geqslant 0, \quad \sum_{j \in X} p_{ij}(s, t) = 1, \quad p_{ij}(s, s) = \delta_{ij}.$$

Soit $f(j)$ fonction arbitraire sur X. L'opérateur T_{st} agit ici selon la formule

$$f_{st}(i) = T_{st} f(i) = \sum_{j \in X} p_{ij}(s, t) f(j).$$

Si m est une mesure arbitraire sur $\mathfrak{B}$ et $m(j) = m(\{j\})$, l'opérateur T_{ts}^* est défini par

$$m_{ts}(j) = T_{ts}^* m(\{j\}) = \sum_{i \in X} m(i) p_{ij}(s, t).$$

L'équation de Kolmogorov-Chapman s'écrit

$$p_{ij}(s, t) = \sum_{k \in X} p_{ik}(s, u) p_{kj}(u, t) \qquad (s \leqslant u \leqslant t).$$

Les relations précédentes peuvent être abrégées. Supposons que les éléments de l'espace X sont ordonnés. Soient $P(s, t)$ la matrice d'éléments $p_{ij}(s, t)$, f le vecteur colonne d'éléments $f(i)$, m le vecteur de composantes $m(i)$, $P^*(s, t)$ la transposée de $P(s, t)$, m^* le vecteur ligne transposé de la matrice m. On a

$$T_{st} f = P(s, t) f,$$
$$T_{st}^* m = m^* P^*(s, t),$$
$$P(s, t) = P(s, u) P(u, t) \qquad (s \leqslant u \leqslant t).$$

L'ensemble $\mathscr{D}$ est composé de toutes les suites $\{f(j),\ j \in X\}$ pour lesquelles la limite

$$(A_s f)(i) = \lim_{h \downarrow 0} \sum_{j \in X} \frac{p_{ij}(s-h,\ s)-\delta_{ij}}{h} f(j)$$

existe quel que soit $i \in X$, $s \in]\,0,\ t\,[$ et

$$\lim_{h \downarrow 0} \sum_{j \in X} p_{ij}(t-h,\ t) f(j) = f(i).$$

Ici

$$\delta_{ij} = \begin{cases} 1 & \text{pour } i = j, \\ 0 & \text{pour } i \neq j. \end{cases}$$

Si par exemple pour tout couple $(i, j) \in X \times X$ et $s \in]\,0,\ t]$ existe la limite finie

$$a_{ij}(s) = \lim_{h \downarrow 0} \frac{p_{ij}(s-h,\ s)-\delta_{ij}}{h}, \tag{11}$$

alors $\mathscr{D}$ contient toutes les suites $f(j)$ telles que $\sum_{j \in X} |f(j)| < \infty$ et

$$(A_s f)(i) = \sum_{j \in X} a_{ij}(s) f(j). \tag{12}$$

La dernière remarque n'est pas toujours suffisante et doit être renforcée.

A ce propos on notera que si les limites (11) existent, alors

$$-a_i(s) = a_{ii}(s) \leqslant 0, \quad a_{ij}(s) \geqslant 0 \quad (i \neq j).$$

De l'inégalité

$$\frac{1-p_{ii}(s-h,\ s)}{h} \geqslant \sum_{j \in J} \frac{p_{ij}(s-h,\ s)}{h},$$

où J est un ensemble fini d'indices et $i \notin J$, on déduit que

$$\sum_j^{(i)} a_{ij}(s) \leqslant a_i(s), \tag{13}$$

où $\sum_j^{(i)}$ indique que la somme est étendue à tous les $j \in X \setminus \{i\}$. Dans les cas suffisamment réguliers, par exemple si la série

$$\sum_j^{(i)} \frac{p_{ij}(s-h,\ s)}{h}$$

converge uniformément en $h > 0$ quel que soit $s \geqslant 0$, l'inégalité (13) peut être remplacée par l'égalité

$$\sum_j^{(i)} a_{ij}(s) = a_i(s). \tag{14}$$

LEMME 1. *Si, quels que soient* $(i, j) \in X \times X$ *et* $s > 0$, *existent les limites* (11) *et est réalisée l'égalité* (14), *alors* $\mathcal{D}$ *contient toutes les suites bornées* $\{f(j),\ j \in X\}$.

Démonstration. On remarquera que par hypothèse la série (12) est absolument convergente. Sans restreindre la généralité on peut admettre que $\sup |f(j)| \leqslant 1$. Considérons la différence

$$\Delta = \sum_{j \in X} \frac{p_{ij}(s-h, s) - \delta_{ij}}{h} f(j) - \sum_{j \in X} a_{ij}(s) f(j).$$

Prenons un nombre arbitraire $\varepsilon > 0$ et cherchons un ensemble fini $J \subset X$ tel que $i \in J$ et

$$\sum_{j \in X \setminus J} a_{ij}(s) = -\sum_{j \in J} a_{ij}(s) < \frac{\varepsilon}{4}.$$

On a désormais

$$|\Delta| \leqslant \left| \sum_{j \in J} \left[\frac{p_{ij}(s-h, s) - \delta_{ij}}{h} - a_{ij} \right] \right| + \sum_{j \in X \setminus J} \frac{p_{ij}(s-h, s)}{h} + \frac{\varepsilon}{4}.$$

Cherchons un $h_0 > 0$ tel que le premier terme de droite soit inférieur à $\frac{\varepsilon}{4}$ pour tous les $h \in]0, h_0[$. On constate que

$$\sum_{j \in X \setminus J} \frac{p_{ij}(s-h, s)}{h} = \sum_{j \in J} \left(\frac{\delta_{ij} - p_{ij}(s-h, s)}{h} + a_{ij}(s) \right) - \sum_{j \in J} a_{ij}(s) < \frac{\varepsilon}{2}.$$

Donc $|\Delta| < \varepsilon$ lorsque $h \in]0, h_0[$.

Pour déduire la deuxième équation de Kolmogorov pour les processus considérés renforçons la condition d'existence des limites (11) de la condition d'existence des limites

$$\lim_{s_1 \uparrow s,\ s_2 \downarrow s} \frac{p_{ij}(s_1, s_2) - \delta_{ij}}{s_2 - s_1} = a_{ij}(s). \tag{15}$$

A noter qu'on démontre par analogie au lemme 1 la proposition suivante: si en un point s, pour tous les j, existent les limites (15) et est réalisée l'égalité (14), la série

$$\sum_{j \in X} \frac{p_{ij}(s_1, s_2) - \delta_{ij}}{s_2 - s_1} \tag{16}$$

est uniformément convergente en s_1 et s_2, $s_1 < s < s_2$, $s_2 - s_1 < h_0$.

THÉORÈME 1. *Si, quels que soient* $(i, j, s) \in X \times X \times]0, t[$, *les limites* (15) *existent et vérifient les égalités* (14), *les probabilités* $p_{ij}(s, t)$ *sont dérivables par rapport à* s, $0 < s < t$, *et sont solutions du système d'équations différentielles (premier système d'équations de*

Kolmogorov)

$$-\frac{\partial p_{ij}(s, t)}{\partial s} = \sum_{x \in X} a_{ik}(s) p_{kj}(s, t). \tag{17}$$

D é m o n s t r a t i o n. Comme

$$-\frac{\partial p_{ij}(s, t)}{\partial s} = \lim_{s_1 \uparrow s,\ s_2 \downarrow s} \frac{p_{ij}(s_1, t) - p_{ij}(s_2, t)}{s_2 - s_1} =$$
$$= \lim_{s_1 \uparrow s,\ s_2 \downarrow s} \sum_{j \in X} \frac{p_{ik}(s_1, s_2) - \delta_{ik}}{s_2 - s_1} p_{kj}(s_2, t),$$

il vient

$$-\frac{\partial p_{ij}(s, t)}{\partial s} - \sum_{j \in X} a_{ik} p_{kj}(s, t) =$$
$$= \lim_{s_1 \uparrow s,\ s_2 \downarrow s} \sum_{j \in X} \left[\frac{p_{ik}(s_1, s_2) - \delta_{ik}}{s_2 - s_1} - a_{ik}(s)\right] p_{kj}(s_2, t) +$$
$$+ \lim_{s_2 \downarrow s} \sum_{j \in X} a_{ik}(s) [p_{kj}(s_2, t) - p_{kj}(s_1, t)] = 0$$

compte tenu de la convergence uniforme de la série (16). ■

Dans la déduction du deuxième système d'équations de Kolmogorov on se bornera au cas d'un nombre fini d'états (X est composé d'un nombre fini de points). On supposera l'existence des limites (15). Les relations (14) sont automatiquement réalisées.

Soit

$$m_j(t) = \sum_{k \in X} m_k p_{kj}(s, t), \quad t > s.$$

Alors ($t_1 < t < t_2$)

$$\frac{m_j(t_2) - m_j(t_1)}{t_2 - t_1} = \sum_{k \in X} m_k(t_1) \frac{p_{kj}(t_1, t_2) - \delta_{kj}}{t_2 - t_1} \to \sum_{k \in X} m_k(t) a_{kj}(t).$$

Donc

$$\frac{\partial m_j(t)}{\partial t} = \sum_{k \in X} m_k(t) a_{kj}(t), \quad t > s. \tag{18}$$

En particulier

$$\frac{\partial p_{ij}(s, t)}{\partial t} = \sum_{k \in X} p_{ik}(s, t) a_{kj}(t), \quad t > s. \tag{19}$$

Dans le cas d'un nombre fini d'états les équations (17) ou (19) forment un système d'équations différentielles ordinaires linéaires possédant une solution unique vérifiant les conditions initiales

$p_{kj}(s, s) = \delta_{kj}$ sous des hypothèses assez larges sur la fonction $a_{kj}(s)$. Donc chaque système d'équations de Kolmogorov définit univoquement les probabilités de passage.

Processus discontinus au sens large. Dans les cas assez réguliers un processus markovien à nombre dénombrable d'états est susceptible d'être un modèle de processus de la nature suivante : le point mobile se trouve dans l'état initial pendant un intervalle de temps aléatoire, puis avec des probabilités déterminées il passe à un autre état où il reste pendant un intervalle de temps aléatoire et ainsi de suite. De tels processus peuvent être étudiés dans un espace des phases arbitraire. On les appelle *processus markoviens discontinus.*

Soit un processus markovien au sens large dans un espace des phases $\{X, \mathfrak{B}\}$ avec la probabilité de passage $\mathsf{P}(s, x, t, B)$, $s < t$, $(s, t) \in I \times I$. On supposera que la tribu $\mathfrak{B}$ contient tous les sous-ensembles à un seul point de X.

DÉFINITION. *Un processus markovien au sens large est discontinu si, quels que soient* $(s, x, B) \in I \times X \times \mathfrak{B}$, *existe la limite*

$$\lim_{t \downarrow s} \frac{\mathsf{P}(s, x, t, B) - \chi(B, x)}{t - s} = \bar{a}(s, x, B) \tag{20}$$

et $\bar{a}(s, x, B)$ *sont des charges finies sur* $\mathfrak{B}$ *pour* (s, x) *fixes.*

Un processus markovien discontinu au sens large est par définition *régulier* si dans la formule (20) la convergence est uniforme en $(s, x, B) \in [0, t] \times X \times \mathfrak{B}$ et la fonction $\bar{a}(s, x, B)$ est continue en $s \in [0, t]$ uniformément par rapport à (x, B), où t est un nombre quelconque de I.

La fonction $\bar{a}(s, x, B)$ possède les propriétés suivantes :

$$\bar{a}(s, x, X) = 0,$$

$$\bar{a}(s, x, B) = \lim_{t \downarrow s} \frac{\mathsf{P}(s, x, t, B)}{t - s} \geqslant 0 \quad \text{si} \quad x \bar{\in} B,$$

$$\bar{a}(s, x, \{x\}) = -\bar{a}(s, x, X \setminus \{x\}) = \lim_{t \downarrow s} \frac{\mathsf{P}(s, x, t, \{x\}) - 1}{t - s} \leqslant 0,$$

où $\{x\}$ est l'ensemble composé du seul point x. On peut grouper ces expressions dans une formule unique en posant

$$\bar{a}(s, x, B) = -a(s, x)\chi(B, x) + a(s, x, B),$$

où

$$a(s, x) = -\bar{a}(s, x, \{x\}), \quad a(s, x, B) = \bar{a}(s, x, B \setminus \{x\}),$$

$a(s, x, B)$ étant une mesure finie sur $\mathfrak{B}$ et $a(s, x\,\{x\}) = 0$.

Le fait que $\bar{a}(s, x, B)$ est une charge finie sur $\mathfrak{B}$ pour un processus discontinu régulier découle de la convergence uniforme dans la formule (20).

En effet, de la définition de la fonction $a(s, x, B)$ il résulte immédiatement que cette fonction est une fonction d'ensembles

sur $\mathfrak{B}$ additive non négative. Supposons maintenant que $B_n \subset B_{n+1}$, $B_n \in \mathfrak{B}$, $B = \bigcup_{n=1}^{\infty} B_n$ et $x \bar{\in} B$. Il vient

$$a(s, x, B) = \lim_{t \downarrow s} \frac{\mathsf{P}(s, x, t, B)}{t-s} \lim_{t \downarrow s} \lim_{n\to\infty} \frac{\mathsf{P}(s, x, t, B)}{t-s} =$$

$$= \lim_{n\to\infty} \lim_{t \downarrow s} \frac{\mathsf{P}(s, x, t, B_n)}{t-s} = \lim_{n\to\infty} a(s, x, B_n),$$

où il est possible d'intervertir l'ordre de passage à la limite en raison de la convergence uniforme par rapport à $B \in \mathfrak{B}$ dans la formule (20). On a donc démontré que la fonction $a(s, x, B)$ est dénombrablement additive.

Signalons encore une propriété du processus discontinu régulier. Pour tout $t \in I$ il existe une constante K telle que

$$|\bar{a}(s, x, B)| \leqslant K \text{ pour tous les } (s, x, B) \in [0, t] \times X \times \mathfrak{B}.$$

Dans la suite de ce point, on n'envisagera que des processus discontinus réguliers au sens large. Posons

$$\Pi(t, x, B) = \begin{cases} \dfrac{a(t, x, B)}{a(t, x)} & \text{si } a(t, x) > 0, \\ \chi(B, x) & \text{si } a(t, x) = 0. \end{cases}$$

Si (t, x) sont fixes, $\Pi(t, x, B)$ est une mesure probabiliste sur $\mathfrak{B}$ qui admet une interprétation probabiliste simple. De (20) il suit

$$\mathsf{P}(t, x, t+\Delta t, \{x\}) = 1 - (a(t, x) + \varepsilon)\Delta t,$$

où $\varepsilon \to 0$ avec Δt. Donc aux infiniment petits près d'ordre supérieur $a(t, x)\Delta t$ est la probabilité que le point mobile quitte à l'instant $t + \Delta t$ l'état x pris à l'instant t. Par ailleurs si $a(t, x) \neq 0$, on a

$$\Pi(t, x, B) = \lim_{\Delta t\to 0} \frac{\mathsf{P}(t, x, t+\Delta t, B \setminus \{x\})}{\mathsf{P}(t, x, t+\Delta t, X \setminus \{x\})},$$

si bien que $\Pi(t, x, B)$ peut être interprétée comme la probabilité conditionnelle que le système quitte à l'instant t l'état x et effectue un saut qui l'amène dans l'ensemble B. Cette interprétation de la fonction $\Pi(t, x, B)$ sera justifiée ultérieurement (cf. § 1, chapitre VII).

La relation (20) s'écrit encore

$$\mathsf{P}(s, x, t, B) = (1 - a(s, x))(t-s)\chi(B, x) + \\ + (a(s, x, B) + r(s, x, t, B))(t-s), \quad (21)$$

où $r(s, x, t, B)$ est une fonction qui tend uniformément en (s, x, B), $s \in [0, t]$, vers zéro lorsque $t \downarrow s$.

La dernière relation entraîne en particulier que

$$\forall u \in I, \quad |\mathsf{P}(s, x, t, B) - \chi(B, x)| \leqslant K_1(t-s), \quad (22)$$

où K_1 est une constante indépendante de $(s, x, t, B) \in [0, u] \times \times X \times [0, u] \times \mathfrak{B}$.

Déduisons maintenant les équations de Kolmogorov pour les processus discontinus. Commençons par la deuxième équation.

Soient s fixe, $t > s$, m mesure probabiliste arbitraire sur $\mathfrak{B}$ et $m_t(B) = T^*_{ts}m(B)$, où T^*_{ts} est l'opérateur introduit plus haut. Si $t_2 > t_1 > s$, on a

$$m_{t_2}(B) - m_{t_1}(B) = \int_X m_{t_1}(dx)\,[\mathsf{P}(t_1, x, t_2, B) - \chi(B, x)],$$

d'où, en vertu de l'inégalité (22), il vient

$$\sup_B |m_{t_2}(B) - m_{t_1}(B)| \leqslant K_1(t_2 - t_1).$$

Par ailleurs (21) entraîne

$$m_{t_2}(B) - m_{t_1}(B) = (t_2 - t_1)\int_X [\bar{a}(t_1, x, B) + r(t_1, x, t_2, B)]\,m_{t_1}(dx). \quad (23)$$

Supposons maintenant que $t_2 \downarrow t$, $t_1 \uparrow t$. Il vient

$$\sup_{(x,B)}\left|\frac{m_{t_2}(B) - m_{t_1}(B)}{t_2 - t_1} - \int_X \bar{a}(t, x, B)\,m_t(dx)\right| \leqslant \sup_{(x,B)} |r(t_1, x, t_2, B)| +$$

$$+ \sup_{(x,B)}\left|\int_X [\bar{a}(t_1, x, B) - \bar{a}(t, x, B)]\,m_{t_1}(dx)\right| +$$

$$+ \sup_{(x,B)}\left|\int_X \bar{a}(t, x, B)\,[m_{t_1}(dx) - m_t(dx)]\right| \leqslant$$

$$\leqslant \varepsilon_1 + \varepsilon_2 + K \sup_B |m_{t_1}(B) - m_t(B)|,$$

où

$$\varepsilon_1 = \sup_{(x,B)} |r(t_1, x, t_2, B)| \to 0 \qquad \text{si } t_1 \uparrow t,\ t_2 \downarrow t,$$

$$\varepsilon_2 = \sup_{(x,B)} |\bar{a}(t_1, x, B) - \bar{a}(t, x, B)| \to 0 \qquad \text{si } t_1 \uparrow t.$$

De (23) il résulte que $\sup_B |m_{t_1}(B) - m_t(B)| \to 0$ lorsque $t_1 \uparrow t$. Ce qui démontre le

THÉORÈME 2. *Dans le cas d'un processus discontinu régulier les fonctions $T^*_{ts}m(B)$ de t sont dérivables pour $t > s$ et*

$$\frac{dT^*_{ts}m}{dt} = A^*_s T^*_{ts} m, \quad (24)$$

où

$$A_s^* m(B) = \int_X \bar{a}(t, y, B)\, m(dy) = -\int_B a(t, y)\, m(dy) + \\ + \int_X a(t, y, B)\, m(dy).$$

La formule (24) *s'écrit encore*

$$\frac{dm_t(B)}{dt} = -\int_B a(t, y)\, m_t(dy) + \int_X a(t, y, B)\, m_t(dy). \qquad (25)$$

En posant $m(B) = \chi(B, x)$ on obtient $m_t(B) = \mathsf{P}(s, x, t, B)$ et le théorème 2 entraîne le

COROLLAIRE. *Les probabilités de passage* $\mathsf{P}(s, x, t, B)$ *d'un processus discontinu régulier sont dérivables par rapport à* t *pour* $t > x$ *et*

$$\frac{\partial \mathsf{P}(s, x, t, B)}{\partial t} = \\ = -\int_B a(t, y)\, \mathsf{P}(s, x, t, dy) + \int_X a(t, y, B)\, \mathsf{P}(s, x, t, dy). \qquad (26)$$

La relation (20) donne la condition initiale suivante pour les équations différentielles (25) et (26) :

$$\lim_{t \downarrow s} m_t(B) = m(B), \quad \lim_{t \downarrow s} \mathsf{P}(s, x, t, B) = \chi(B, x). \qquad (27)$$

La solution (si elle existe) de l'équation (26) qui vérifie cette condition initiale définit les probabilités de passage du processus envisagé.

Déduisons maintenant la première équation de Kolmogorov. Soit t fixe, $f(x) \in \mathscr{B}(\mathfrak{B})$, $\|f\| = \sup_x |f(x)|$ et

$$f_s(x) = T_{st} f(x) = \int_X f(y)\, \mathsf{P}(s, x, t, dy), \quad s < t.$$

Posons $s_1 < s < s_2 < t$. Il vient

$$f_{s_2}(x) - f_{s_1}(x) = \int [f_{s_2}(x) - f_{s_2}(y)]\, \mathsf{P}(s_1, x, s_2, dy) = \\ = (s_2 - s_1) \int [f_{s_2}(x) - f_{s_2}(y)]\, (\bar{a}(s_1, x, dy) + r(s_1, x, s_2, dy)).$$

Cette relation entraîne

$$\sup_x |f_{s_2}(x) - f_{s_1}(x)| \leqslant 2\|f\|(s_2 - s_1)\, |[K + 2 \sup_{(x, B)} |r(s_1, x, s_2, B)|].$$

Par ailleurs, eu égard à $\overline{a}(s, x, X) = 0$ on obtient

$$\left\| \frac{f_{s_2}(x) - f_{s_1}(x)}{s_2 - s_1} + \int f_s(y) \overline{a}(s, x, dy) \right\| \leqslant$$

$$\leqslant \left\| \int [f_s(y) - f_{s_2}(y)] \overline{a}(s, x, dy) \right\| +$$

$$+ \left\| \int f_{s_2}(y) [\overline{a}(s, x, dy) - \overline{a}(s_1, x, dy)] \right\| +$$

$$+ \left\| \int [f_{s_2}(x) - f_{s_2}(y)] r(s_1, x, s_2, dy) \right\|. \quad (28)$$

En vertu de (28), le second membre de l'inégalité n'excède pas

$$2 \| f \| (s_2 - s) [K + 2 \sup_{(x, B)} | r(s_1, x, s_2, B) |] +$$

$$+ \| f \| 2 \sup_{(x, B)} | \overline{a}(s, x, B) - \overline{a}(s_1, x, B) | +$$

$$+ 2 \| f \| \cdot 2 \sup_{(x, B)} | r(s_1, x, s_2, B) |.$$

La régularité du processus discontinu entraîne que l'expression obtenue tend vers zéro pour $s_1 \uparrow s$ et $s_2 \downarrow s$. Ce qui démontre le

THÉORÈME 3. *Dans le cas d'un processus discontinu régulier au sens large la fonction $f_s(x) = T_{st} f(x)$, $s < t$, est dérivable par rapport à s (uniformément en x) et est solution de l'équation*

$$\frac{\partial f_s(x)}{\partial s} = - \int f_s(y) \overline{a}(s, x, dy) =$$

$$= a(s, x) \left[f_s(x) - \int f_s(y) \Pi(s, x, dy) \right], \quad s < t, \quad (29)$$

avec la condition aux limites

$$\lim_{s \uparrow t} f_s(x) = f(x). \quad (30)$$

L'équation (29) est la première équation de Kolmogorov pour les processus discontinus réguliers; l'opérateur A_s s'écrit ici

$$A_s f(x) = a(s, x) \left[-f(x) + \int f(y) \Pi(s, x, dy) \right].$$

COROLLAIRE. *Les probabilités de passage $\mathsf{P}(s, x, t, B)$ d'un processus discontinu régulier sont dérivables par rapport à s, $s < t$, et sont solutions de l'équation*

$$\frac{\partial \mathsf{P}(s, x, t, B)}{\partial s} =$$

$$= a(s, x) \left[\mathsf{P}(s, x, t, B) - \int \mathsf{P}(s, y, t, B) \Pi(s, x, dy) \right] \quad (31)$$

avec la condition aux limites

$$\lim_{s \uparrow t} \mathsf{P}(s, x, t, B) = \chi(B, x).$$

Montrons maintenant que sous certaines conditions les fonctions $a(t, x)$ et $a(t, x, B)$ définissent de façon unique un processus markovien au sens large discontinu régulier. Il s'agit en premier lieu des solutions des équations (26) ou (29) vérifiant des conditions aux limites appropriées.

Soit $I = [0, t^*[$. Le processus discontinu étant régulier, on imposera aux fonctions $a(t, x)$ et $a(t, x, B)$ les conditions suivantes:

a) pour $(t, x) \in I \times X$ fixes la fonction $a(t, x, B)$ est une mesure sur $\mathfrak{B}$, $a(t, x, \{x\}) = 0$ et $a(t, x) = a(t, x, X)$;

b) pour (x, B) fixes la fonction $a(t, x, B)$, $t \in I$, est continue en t uniformément en (x, B) et pour (t, B) fixes $\mathfrak{B}$-mesurable comme fonction de x.

Considérons l'espace $\mathscr{W} = \mathscr{W}(\mathfrak{B})$ des fonctions $w(B)$ complètement additives finies (charges finies) définies sur l'espace mesurable $\{X, \mathfrak{B}\}$. Définissons dans $\mathscr{W}$ une distance $\rho(w_1, w_2)$ telle que

$$\rho(w_1, w_2) = \| w_1(B) - w_2(B) \|, \quad w_i \in \mathscr{W},$$

où

$$\| w(B) \| = \sup \{ | w(B) |, \ B \in \mathfrak{B} \}.$$

On s'assure aisément que $\mathscr{W}$ est un espace vectoriel noımé complet. On comprendra les équations (25) et (27) comme des équations définies sur $\mathscr{W}$ et l'on interprétera en conséquence la notion de dérivée dans le second membre de (25).

Considérons encore l'espace $\mathscr{C}^{\mathscr{W}}[s, t^*]$ des fonctions continues $\check{w} = \check{w}_t = w_t(B)$, $t \in [s, t^*]$ à valeurs dans $\mathscr{W}$ et de norme $|||\check{w}||| = \max \{ \| \check{w}_t \|, \ t \in [s, t^*] \}$.

THÉORÈME 4. *Si la fonction $a(t, x, B)$ réalise les conditions* a) *et* b), *le système d'équations* (25) *et* (27) *admet une solution unique dans* $\mathscr{C}^{\mathscr{W}}$. *Cette solution est une mesure si $m(B)$ l'est.*

D é m o n s t r a t i o n. On remarquera que la condition b) entraîne que la fonction $a(t, x)$ est uniformément bornée en (t, x), $a(t, x) \leqslant K < \infty$. Soit dans $\mathscr{C}^{\mathscr{W}}[s, t^*]$ la fonction

$$q_t(B) = \int_B \exp \left\{ \int_s^t a(\theta, x) \, d\theta \right\} m_t(dx).$$

Si la fonction m_t est dérivable dans $\mathscr{W}$, q_t le sera également, et

inversement, et de plus

$$\frac{dq_t(B)}{dt}=\int_B a(t,\ x)\,q_t(dx)+\int_B \exp\left\{\int_s^t a(\theta,\ x)\,d\theta\right\}\frac{dm_t}{dt}(dx).$$

En remplaçant $\frac{dm_t}{dt}$ par son expression tirée de (25) on obtient

$$\frac{dq_t(B)}{dt}=\int_B\int_X \exp\left\{\int [a(\theta,\ x)-a(\theta,\ y)]\,d\theta\right\}a(t,\ y,\ dx)\,q_t(dy)=$$

$$=\int_X b(t,\ y,\ B)\,q_t(dy),$$

où

$$b(t,\ y,\ B)=\int_B \exp\left\{\int_s^t [a(\theta,\ x)-a(\theta,\ y)]\,d\theta\right\}a(t,\ y,\ dx).$$

Donc les équations (25) et (27) sont équivalentes à l'équation

$$q_t(B)=m(B)+\int_s^t\int_X b(\theta,\ y,\ B)\,q_\theta(dy)\,d\theta,\quad t\in[s,\ t^*],\qquad (32)$$

où $b\ (t,\ y,\ B)$ est uniformément bornée, $b\ (t,\ y,\ B)\leqslant K_1$ et comme fonction de B est une mesure. L'opérateur Q^* dans $\mathscr{C}^{\mathscr{W}}$ défini par

$$(Q^*\check{w})_t\,(B)=m\,(B)+\int_s^t\int_X b\,(\theta,\ y,\ B)\,w_\theta\,(dy)\,d\theta$$

vérifie les relations

$$\|(Q^*\check{w}')_t-(Q^*\check{w}'')_t\|\leqslant 2K_1(t-s)\,|||\check{w}'-\check{w}''|||,$$

$$\|(Q^{*n}\check{w}')_t-(Q^{*n}\check{w}'')_t\|\leqslant (2K_1)^n\frac{(t-s)^n}{n!}\,|||\check{w}'-\check{w}''|||,$$

où Q^{*n} désigne la n-ième puissance de Q^*. Donc une puissance de Q^* est un opérateur contractant. En vertu du principe des applications contractantes, l'équation (32) admet une solution unique dans $\mathscr{C}^{\mathscr{W}}$ qui peut être obtenue par la méthode des approximations successives; donc si $m\ (B)$ est une mesure, $q_t\ (B)$ le sera également. ■

On étudie l'équation (29) de façon analogue. La substitution

$$f_s(x)=\exp\left\{-\int_s^t a(\theta,\ x)\,d\theta\right\}g_s(x)$$

ramène l'équation (29) à l'équation équivalente plus simple

$$\frac{\partial g_s(x)}{\partial t} = -\int_X g_s(y) \exp\left\{\int_s^t [a(\theta, x) - a(\theta, y)]\, d\theta\right\} a(s, x, dy),$$

($s < t$) avec la condition aux limites $g_t(x) = f(x)$. Cette équation équivaut à son tour à

$$g_s(x) = f(x) + \\ + \int_s^t \int_X g_v(y) \exp\left\{\int [a(\theta, x) - a(\theta, y)]\, d\theta\right\} a(v, x, dy)\, dv. \quad (33)$$

Introduisons l'espace $\mathscr{C}^{\mathscr{B}(\mathfrak{B})}[0, t]$ des fonctions continues $\check{f} = f_s = f_s(x)$ de s à valeurs dans $\mathscr{B}(\mathfrak{B})$ et de norme $|||\ \ ||| = \sup\{|f_s(x)|, (s, x) \in [0, t] \times X\}$. Dans $\mathscr{C}^{\mathscr{B}(\mathfrak{B})}[0, t]$ l'opérateur linéaire Q:

$$(Qg)_s(x) = f(x) + \\ + \int_s^t \int_X g_v(y) \exp\left\{\int_v^t [a(\theta, x) - a(\theta, y)]\, d\theta\right\} a(v, x, dy)\, dv,$$

applique l'ensemble des fonctions non négatives de $\mathscr{C}^{\mathscr{B}(\mathfrak{B})}[0, t]$ dans lui-même et de plus

$$\|(Q\check{g}')_s - (Q\check{g}'')_s\| \leqslant K_2(t-s)\,|||\check{g}' - \check{g}''|||,$$
$$\|(Q^n\check{g}')_s - (Q^n\check{g}'')_s\| \leqslant \frac{K_2^n(t-s)^n}{n!}\,|||\check{g}' - \check{g}''|||,$$

où $K_2 = Ke^{Kt}$. Donc une puissance de Q est opérateur contractant et l'équation (33), et, par voie de conséquence, l'équation (29), possède dans $\mathscr{C}^{\mathscr{B}(\mathfrak{B})}[0, t]$ une solution unique vérifiant la condition aux limites (30).

THÉORÈME 5. *Si la fonction $a(t, x, B)$ vérifie les conditions* a) *et* b), *les équations* (29), (30) *admettent une solution unique. En particulier, dans le cas considéré les probabilités de passage affectées au processus sont définies de façon unique par la fonction $a(t, x, B)$.*

REMARQUE. L'équation (33) peut être résolue par la méthode des approximations successives. A cet effet on peut mettre la solution des équations (29), (30) sous la forme

$$f_s(x) = \sum_{n=0}^{\infty} f_s^{(n)}(x),$$

où

$$f_s^{(0)}(x)=\exp\left\{-\int_s^t a(\theta,\ x)\,d\theta\right\} f(x),$$

$$f_s^{(n+1)}(x)=\int_s^t \int_X f_v^{(n)}(y)\exp\left\{-\int_s^v a(\theta,\ x)\,d\theta\right\} a(v,\ x,\ dy)\,dv.$$

En particulier, pour les probabilités de passage $\mathsf{P}(s, x, t, B)$ on obtient les expressions suivantes:

$$\mathsf{P}(s,\ x,\ t,\ B)=\sum_{n=0}^{\infty} \mathsf{P}^{(n)}(s,\ x,\ t,\ B), \tag{34}$$

où

$$\mathsf{P}^{(0)}(s,\ x,\ t,\ B)=\exp\left\{-\int_s^t a(\theta,\ x)\,d\theta\right\}\chi(B,\ x), \tag{35}$$

$$\mathsf{P}^{(n+1)}(s,\ x,\ t,\ B)=$$

$$=\int_s^t \int_X \mathsf{P}^{(n)}(v,\ y,\ t,\ B)\exp\left\{-\int_s^v a(\theta,\ x)\,d\theta\right\} a(v,\ x,\ dy)\,dv. \tag{36}$$

Les fonctions $\mathsf{P}^{(n)}(s, x, t, B)$ admettent une interprétation probabiliste simple qui sera donnée au § 2 du chapitre VII où l'on montrera également comment connaissant la fonction $a(s, x, B)$ on peut construire un processus markovien sous des conditions plus larges que celles envisagées ici.

Les résultats obtenus peuvent être appliqués aux processus à nombre dénombrable d'états. Dans ce cas l'espace X est composé d'un nombre dénombrable de points et il suffit de considérer les probabilités de passage dans des espaces composés d'un seul point. Supposons que $p_{ij}(s, t) = \mathsf{P}(s,\ i,\ t,\ \{j\})$, $i, j \in X$. Au lieu de $a(s, x, B)$ étudions la fonction

$$a(s,\ i,\ j)=\lim_{t\downarrow s}\frac{p_{ij}(s,\ t)}{t-s},\quad i\neq j,$$

qui est confondue avec la fonction $a_{ij}(s)$. Les conditions a) et b) deviennent ici:

a) $a(t,\ i)=\sum_{j\in X} a(t,\ i,\ j)$, où $a(s,\ i)=\lim_{t\downarrow s}(1-p_{ii}(s,\ t))(t-s)^{-1}$;

b) $a(t, i, j)$ sont continues en t sur $[0, t^*]$ uniformément en (i, j). Si ces conditions sont satisfaites, la première et la deuxième équation de Kolmogorov pour les processus markoviens à nombre

dénombrable d'états possèdent des solutions uniques qui peuvent être obtenues avec les formules indiquées plus haut. Par exemple

$$p_{ij}(s, t) = \sum_{n=0}^{\infty} p_{ij}^{(n)}(s, t),$$

où

$$p_{ij}^{(0)}(s, t) = \exp\left\{-\int_s^t a_i(\theta)\,d\theta\right\}\delta_{ij},$$

$$p_{ij}^{(n+1)}(s, t) = \int_s^t \sum_{k\in X} p_{kj}^{(n)}(v, t)\exp\left\{-\int_s^v a_i(\theta)\,d\theta\right\} a_{ik}(v)\,dv,$$

$$n = 0, 1, 2, \ldots$$

Processus à accroissements indépendants. Ces processus sont un cas particulier des processus markoviens. Soit X espace vectoriel métrique, $\mathfrak{B}$ la tribu des boréliens de X. Désignons par $B + x$ ($B \subset X$, $x \in X$) la x-translation de B: $B + x\ \{y : y = z + x,\ z \in \in B\}$.

Considérons la famille de mesures probabilistes $\mathsf{P}_{st}(\cdot)$ sur $\mathfrak{B}$ ($s \geqslant 0$, $t > s$) vérifiant les conditions suivantes:

a) $\mathsf{P}_{st}(B - x)$ est une fonction $\mathfrak{B}$-mesurable de x quel que soit $B \in \mathfrak{B}$;

b) si $s < u < t$, on a

$$\mathsf{P}_{st}(B) = \int_X \mathsf{P}_{ut}(B - y)\,\mathsf{P}_{su}(dy). \tag{37}$$

On vérifie immédiatement que pour toute fonction $f(x)$ $\mathfrak{B}$-mesurable on a

$$\int_X f(x+y)\,\mathsf{P}_{st}(dy) = \int_X f(y)\,\mathsf{P}_{st}(dy - x)$$

(cette relation est triviale pour les indicateurs des ensembles $\mathfrak{B}$-mesurables). Donc (37) entraîne

$$\mathsf{P}_{st}(B - x) = \int_X \mathsf{P}_{ut}(B - y)\,\mathsf{P}_{su}(dy - x).$$

Si donc l'on pose $\mathsf{P}(s, x, t, B) = \mathsf{P}_{st}(B - x)$, la fonction $\mathsf{P}(s, x, t, B)$ sera probabilité de passage. Elle présente une homogénéité spatiale, c'est-à-dire

$$\mathsf{P}(s, x + y, t, B + y) = \mathsf{P}(s, x, t, B) \quad \forall y \in X$$

Inversement, si des probabilités de passage jouissent de cette propriété, alors $\mathsf{P}(s, x, t, B) = \mathsf{P}_{st}(B - x)$.

Définissons sur $\{X, \mathfrak{B}\}$ une mesure probabiliste et considérons la famille de répartitions $\{\mathsf{P}_{t_1, \ldots, t_n}, 0 \leqslant t_1 < \ldots < t_n, n = 1, 2, \ldots\}$, où $\mathsf{P}_{t_1, \ldots, t_n}(B^{(n)})$ est la répartition définie sur $\{X^n, \mathfrak{B}^n\}$ $(B^{(n)} \in \mathfrak{B}^n)$ par

$$\mathsf{P}_{t_1, \ldots, t_n}(B^{(n)}) = \int\limits_X \left\{ \int\limits_{B^{(n)}} \mathsf{P}(0, x_0, t_1, dx_1)\, \mathsf{P}(t_1, x_1, t_2\, dx_2) \ldots \right.$$
$$\left. \ldots \mathsf{P}(t_{n-1}, x_{n-1}, t_n, dx_n) \right\} q(dx_0).$$

On s'assure sans peine que cette famille de répartitions est associée à un processus à accroissements indépendants.

Au § 3 nous avons examiné en détail la structure d'une famille de mesures P_{st} vérifiant (37) dans le cas d'un espace X finidimensionnel ($X = \mathscr{R}^d$) et d'un processus à accroissements indépendants stochastiquement continu et homogène dans le temps (c'est-à-dire $\mathsf{P}_{st}(B) = \mathsf{P}_{t-s}(B)$). Dans ces conditions la fonction caractéristique $\varphi(t, u)$ de la répartition $\mathsf{P}_t(B)$ peut s'écrire

$$\varphi(t, u) = \int\limits_{\mathscr{R}^d} e^{i(u, x)} \mathsf{P}_t(dx) = \exp\left(t \left\{ i(a, u) - \frac{1}{2}(bu, u) + \right.\right.$$
$$\left.\left. + \int\limits_{\mathscr{R}^d} \left[e^{i(u, z)} - 1 - \frac{i(u, z)}{1+|z|^2} \right] \frac{1+|z|^2}{|z|^2} \Pi(dz) \right\}\right), \qquad (38)$$

où $a \in \mathscr{R}^d$ et b est une application symétrique linéaire définie non négative de $\mathscr{R}^d$ dans $\mathscr{R}^d$, Π une mesure finie sur $\mathfrak{B}$ et $\Pi\{0\} = 0$.

Posons

$$f_s(x) = \int\limits_{\mathscr{R}^d} f(y)\, \mathsf{P}(s, x, t, dy) = \int\limits_{\mathscr{R}^d} f(x+y)\, \mathsf{P}_{st}(dy), \quad s < t.$$

Il est évident que si $f(x)$ est une fonction deux fois continûment dérivable, bornée avec ses dérivées partielles première et seconde, $f_s(x)$ le sera également. On a

$$\frac{f_{s-h}(x) - f_s(x)}{h} = \int\limits_{\mathscr{R}^d} [f_s(x+z) - f_s(x)] \frac{1}{h} \mathsf{P}_{(s-h,\, h)}(dz) =$$
$$= (\bar{A}_h, \nabla f_s(x)) - \frac{1}{2}(\bar{B}_h \nabla, \nabla f_s(x)) + \int \left[f(x+z) - f(x) - \right.$$
$$\left. - \frac{(z, \nabla) f_s(x)}{1+|z|^2} \frac{1}{2} \frac{(z, \nabla f_s(x))^2}{1+|z|^2} \right] \frac{1+|z|^2}{|z|^2} \Pi_h(dz),$$

où la mesure $\Pi_h\,(dz)$ est difinie par (10), § 3, et

$$(a,\ \nabla) f_s(x) = \sum_1^d a_k \frac{\partial f_s(x)}{\partial x_k},$$

$$\overline{A}_h = \int\limits_{\mathcal{R}^d} \frac{(z,\ \nabla) f_s(x)}{|z|^2} \Pi_h\,(dz),\ \overline{B}_h = \int\limits_{\mathcal{R}^d} \frac{(z,\ \nabla f_s(x))^2}{|z|^2} \Pi_h\,(dz).$$

Les résultats du § 3 (cf. démonstration du théorème 3) entraînent l'existence de la dérivée $\frac{\partial f_s(x)}{\partial s}$, solution de l'équation

$$\frac{\partial f_s(x)}{\partial s} = (a,\ \nabla) f_s(x) - \frac{1}{2}(b\nabla,\ \nabla) f_s(x) + \\ + \int\limits_{\mathcal{R}^d} \left[f_s(x+z) - f_s(x) - \frac{(z,\ \nabla) f_s(x)}{1+|z|^2} \right] \frac{1+|z|^2}{|z|^2} \Pi\,(dz), \qquad (39)$$

où

$$(b\nabla,\ \nabla) f_s(x) = \sum_{k,\ j=1}^d b_{kj} \frac{\partial^2 f_s(x)}{\partial x_k\, \partial x_j}.$$

La dernière équation s'écrit encore (cf. formule (17), § 3)

$$\frac{\partial f_s(x)}{\partial s} = (a,\ \nabla) f_s(x) + \frac{1}{2}(b\nabla,\ \nabla) f_s(x) + \\ + \int\limits_{\mathcal{R}^d \setminus S_c} [f_s(x+z) - f_s(x)]\, \overline{\Pi}\,(dz) + \\ + \int\limits_{S_c} [f_s(x+z) - f_s(x) - (z,\ \nabla) f_s(x)]\, \overline{\Pi}\,(dz), \qquad (40)$$

où S_c est la sphère de $\mathcal{R}^d$ centrée en l'origine des coordonnées et de rayon c, la mesure $\overline{\Pi}$ n'est pas forcément finie mais comme précédemment $\overline{\Pi}\,\{0\} = 0$ et

$$\int\limits_{S_c} |z|^2\, \Pi\,(dz) < \infty,\ \Pi\,(\mathcal{R}^d \setminus S_c) < \infty.$$

Processus markoviens faiblement dérivables au sens large. Quand on étudie des processus markoviens dans un espace finidimensionnel $\mathcal{R}^d$, il semble naturel de s'intéresser à la classe de processus qui ont la même structure locale que les processus à accroissements indépendants. On peut définir ces processus de la façon suivante.

Soit la fonction caractéristique

$$\varphi(s, x, t, u) = \int_{\mathcal{R}^d} e^{i(u, y)} \mathsf{P}(s, x, t, dy), \quad s < t, \quad \varphi(s, x, s, u) = 1,$$

d'une répartition $\mathsf{P}(s, x, t, B)$.

Un processus markovien au sens large est *faiblement dérivable* par définition si la fonction $\varphi(s, x, t, u)$ est dérivable par rapport à s au point $s = t$, uniformément dans un domaine borné de variation de u, c'est-à-dire si la limite

$$g(t, x, u) = \lim_{s \downarrow t} \frac{J(s, x, t, u) - 1}{t - s}$$

existe uniformément en u pour $|u| \leqslant N$, où N est arbitraire pour tous les $x \in \mathcal{R}^d$, $t \in]0, t^*[$.

Les résultats du théorème 1, § 3, entraînent que si un processus markovien est faiblement dérivable, il existe un vecteur $a(s, x) \in \mathcal{R}^d$, une application symétrique définie non négative $b(s, x) \in \mathcal{R}^d$ dans $\mathcal{R}^d$ et une mesure $q(s, x, B)$ sur $\mathfrak{B}$ tels que toute fonction $f(x)$, $x \in \mathcal{R}^d$, deux fois continûment dérivable, bornée avec ses dérivées partielles première et seconde, vérifie

$$A_s f(x) = \lim_{h \downarrow 0} \frac{T_{s-h, s} f(x) - f(x)}{h} = (a(s, x), \nabla) f(x) +$$
$$+ \frac{1}{2}(b(s, x)\nabla, \nabla) f(x) + \int_{\mathcal{R}^d \setminus S} [f(x+z) - f(x)]\, q(s, x, dz) +$$
$$+ \int_S [f(x+z) - f(x) - (z, \nabla) f(x)]\, q(s, x, dz), \quad (41)$$

par ailleurs $q(s, x, \{0\}) = 0$, $q(s, x, \mathcal{R}^d \setminus S) < \infty$ et

$$\int_S |z|^2 p(s, x, dz) < \infty.$$

En particulier, si $q(s, x) = q(s, x, \mathcal{R}^d) < \infty$, la formule précédente devient

$$A_s f(x) = (\tilde{a}(s, x), \nabla) f(x) + \frac{1}{2}(b(s, x)\nabla, \nabla) f(x) -$$
$$-\Big[q(s, x) f(x) - \int_{\mathcal{R}^d} f(x+z)\, q(s, x, dz)\Big]. \quad (42)$$

Si $\tilde{a}(s, x) \equiv 0$, $b(s, x) \equiv 0$, le processus markovien correspondant est, ainsi qu'il résulte de ce qui précède, un processus discontinu.

Dans le cas général la relation (41) admet l'interprétation suivante. Désignons par $\xi(t)$ l'état du système caractérisé par le processus markovien étudié. Supposons que $\xi(s) = x$. Alors $\Delta\xi(s) =$

$= \xi(s + \Delta s) - x$ peut aux infiniment petits d'ordre supérieur près être mis sous la forme $\Delta\xi(s) = \Delta\xi_1 + \Delta\xi_2 + \Delta\xi_3$, où $\Delta\xi_1$ correspond à la composante non aléatoire du déplacement $\Delta\xi(s)$ et peut s'écrire $\Delta\xi_1 = \tilde{a}(s, x)\,\Delta s$, $\Delta\xi_2$ correspond au déplacement d'un processus wienérien de matrice de variance $b(s, x)\,\Delta s$, et $\Delta\xi_3 = 0$ avec la probabilité $1 - q(s, x)\,\Delta s$ et avec la probabilité $q(s, x)$ coïncide avec un vecteur aléatoire de répartition $\frac{q(s, x, B)}{q(s, x)}$. De plus $\Delta\xi_2$ et $\Delta\xi_3$ sont indépendants.

Si $q(s, x, B) \equiv 0$ dans la formule (41), le processus markovien correspondant est un *processus de diffusion*. Dans ce cas la partie principale du déplacement $\Delta\xi(s)$ est composée d'un terme non aléatoire $a(s, \xi_s)\,\Delta s$ (ou vecteur de translation) et d'un terme aléatoire admettant une répartition gaussienne d-dimensionnelle, d'espérance mathématique nulle et de matrice de corrélation $b(s, \xi_s)\,\Delta s$.

Les processus de diffusion jouent un rôle important dans la théorie et les applications des processus markoviens. Ils feront l'objet d'une étude détaillée au chapitre VIII. On se bornera ici à donner une autre définition du processus de diffusion et à préciser la déduction pour lui des équations différentielles de Kolmogorov.

Définition. *Un processus markovien au sens large est un processus de diffusion si*

1) *quel que soit* $x \in \mathcal{R}^d$, $\varepsilon > 0$,

$$\mathsf{P}(s, x, t, \overline{S_\varepsilon \setminus (x)}) = o(t - s) \tag{43}$$

uniformément en $t < s$, *où* $\overline{S_\varepsilon(x)}$ *est le complémentaire de la boule* $S_\varepsilon(x)$ *de rayon* ε *et de centre* x;

2) *existent une fonction* $a(s, x)$ *à valeurs dans* $\mathcal{R}^d$ *et un opérateur symétrique linéaire défini non négatif* $b(s, x)$ *de* $\mathcal{R}^d$ *dans* $\mathcal{R}^d$ $((s, x) \in \in [0, t^*] \times \mathcal{R}^d)$ *tels que pour tout* $x \in X$ *et* $\varepsilon > 0$

$$\int_{S_\varepsilon(x)} (y - x)\,\mathsf{P}(s, x, t, dy) = a(s, x)(t - s) + o(t - s), \tag{44}$$

$$\int_{S_\varepsilon(x)} (z, y - x)^2\,\mathsf{P}(s, x, t, dy) = (b(s, x) z, z)(t - s) + o(t - s) \tag{45}$$

uniformément en s, $s < t$. *Le vecteur* $a(s, x)$ *s'appelle vecteur de translation, l'opérateur* $b(s, x)$ *opérateur de diffusion du processus markovien.*

Soit une base de $\mathcal{R}^d$. Désignons par $a_i(s, x)$, $i = 1, \ldots, m$, les composantes du vecteur $a(s, x)$ et par $b_{ij}(s, x)$, $i, j = 1, \ldots, m$, les éléments de la matrice de l'opérateur $b(s, x)$ par rapport à cette base.

THÉORÈME 6. *Si les fonctions* $a(s, x)$ *et* $b(s, x)$ *sont continues et* $f(x)$ *continue, bornée et telle que la fonction*

$$u(s, x) = \int_{\mathcal{R}^d} f(y)\, \mathrm{P}(s, x, t, dy)$$

possède des dérivées partielles $\frac{\partial u(s, x)}{\partial x_i}$, $\frac{\partial^2 u(s, x)}{\partial x_i \partial x_j}$ *continues, alors* $u(s, x)$ *est dérivable par rapport à* s *et est solution de l'équation*

$$-\frac{\partial u(s, x)}{\partial s} = (a(s, x), \nabla)\, u(s, x) + \frac{1}{2}(b(s, x)\nabla, \nabla)\, u(s, x) \quad (46)$$

avec la condition aux limites

$$\lim_{s \uparrow t} u(s, x) = f(x). \quad (47)$$

Démonstration. Soit $s_1 \leqslant s \leqslant s_2 < t$. La fonction $u(s, x)$ étant bornée, on a

$$u(s_1, x) - u(s_2, x) = \int_{\mathcal{R}^d} [u(s_2, y) - u(s_2, x)]\, \mathrm{P}(s_1, x, s_2, dy) =$$

$$= \int_{S_\varepsilon(x)} [u(s_2, y) - u(s_2, x)]\, \mathrm{P}(s_1, x, s_2, dy) + o_\varepsilon(s_2 - s_1),$$

où $\frac{o_\varepsilon(s_2 - s_1)}{s_2 - s_1} \to 0$ quel que soit $\varepsilon > 0$ fixe. Soit le développement taylorien

$$u(s_2, y) - u(s_2, x) =$$

$$= (y - x, \nabla)\, u(s_2, x) + \frac{1}{2}(y - x, \nabla)^2\, u(s_2, x) + r(x, y, s_2),$$

en outre, lorsque $y \in S_\varepsilon(x)$, on a $|r(x, y, s_2)| \leqslant |y - x|^2 \omega_\varepsilon$, où

$$\omega_\varepsilon = \sup_{i, j, s_2, y \in S_\varepsilon(x)} \left| \frac{\partial^2 (u)(s_2, x + \theta(y - x))}{\partial x_i \partial x_j} - \frac{\partial^2 u(s_2, x)}{\partial x_i \partial x_j} \right| \to 0$$

lorsque $\varepsilon \to 0$. D'où il suit

$$u(s_1, x) - u(s_2, x) =$$

$$= \left[(a(s_2, x), \nabla)\, u(s_2, x) + \frac{1}{2}(b(s_2, x)\nabla, \nabla)\, u(s_2, x) + R' \right](s_2 - s_1), \quad (48)$$

où $\lim_{\varepsilon \to 0} \lim_{s_2 - s_1 \downarrow 0} R' = 0$. En divisant l'expression obtenue par $s_2 - s_1$, en passant à la limite pour $s_2 \downarrow s$, $s_1 \uparrow s$ et du fait de la continuité du second membre de (48) en s_2, on obtient l'équation (46).

La condition aux limites (47) découle de l'égalité

$$u(s, x) - f(x) = \int_{S_\varepsilon(x)} [f(y) - f(x)]\, \mathrm{P}(s, x, t, dy) + o_\varepsilon(t - s),$$

où $\varepsilon > 0$ est un nombre arbitrairement petit, et de la continuité de $f(x)$. ■

Supposons maintenant qu'existe la densité de probabilité de passage, c'est-à-dire une fonction $p(s, x, t, y)$ telle que pour tout ensemble borélien B

$$\mathsf{P}(s, x, t, B) = \int_B p(s, x, t, y)\,dy,$$

où l'intégration est effectuée par rapport à la mesure de Lebesgue dans $\mathscr{R}^d$. L'équation de Kolmogorov-Chapman s'écrit alors

$$p(s, x, t, y) = \int_{\mathscr{R}^d} p(s, x, u, z)\,p(u, z, t, y)\,dz, \quad s < u < t. \tag{49}$$

Si $p(s, x, t, y)$ comme fonction de (t, y) est suffisamment de fois dérivable, elle est solution de la deuxième équation de Kolmogorov qui porte encore le nom d'*équation de Fokker-Planck.*

THÉORÈME 7. *Si les relations* (43), (44), (45) *sont réalisées uniformément en x et existent les dérivées continues*

$$\frac{\partial p(s, x, t, y)}{\partial t}, \quad \frac{\partial (a_i(t, y)\,p(s, x, t, y))}{\partial x_i}, \quad \frac{\partial^2 (b_{ij}(t, y)\,p(s, x, t, y))}{\partial x_i\,\partial x_j},$$

la fonction $p(s, x, t, y)$ est solution de l'équation

$$\frac{\partial p(s, x, t, y)}{\partial t} = -(\nabla, a(t, y)\,p(s, x, t, y)) + \\ + \frac{1}{2}(\nabla, \nabla b(t, y)\,p(s, x, t, y)) \tag{50}$$

pour $(t, y) \in (s, t^*) \times \mathscr{R}^d$.

Démonstration. Soit $g(x)$ une fonction quelconque deux fois continûment dérivable s'annulant en dehors d'un compact. Comme dans la démonstration du théorème précédent, on s'assure que

$$\lim_{t_2 \downarrow t,\, t_1 \uparrow t} \frac{1}{t_2 - t_1}\left[\int g(y)\,p(t_1, x, t_2, y)\,dy - g(x)\right] = \\ = (a(t, x), \nabla)\,g(x) + \frac{1}{2}(b(t, x)\nabla, \nabla)\,g(x)$$

uniformément en x. En se servant des conditions du théorème et de la dernière égalité, on obtient

$$\frac{\partial}{\partial t}\int p(s, x, t, y)\,g(y)\,dy = \\ = \lim_{t_1 \uparrow t,\, t_2 \downarrow t} \frac{1}{t_2 - t_1}\int \left[p(s, x, t_2, y) - p(s, x, t_1, y)\right] g(y)\,dy =$$

$$= \lim_{t_1 \uparrow t,\ t_2 \downarrow t} \int p(s, x, t_1, y) \left[\frac{1}{t_2 - t_1} \int_{\mathcal{R}^d} p(t_1, y, t_2, z)\, g(z)\, dz - g(y) \right] dy =$$

$$= \int p(s, x, t, y) \left[(a(t, y), \nabla)\, g(y) + \frac{1}{2} (b(t, y) \nabla, \nabla)\, g(y) \right] dy.$$

Une intégration par parties donne

$$\int \frac{\partial}{\partial t} p(s, x, t, y)\, g(y)\, dy = - \int \Big[(\nabla, p(s, x, t, y)\, a(t, y)) +$$
$$+ \frac{1}{2} (\nabla, \nabla p(s, x, t, y)\, b(t, y)) \Big] g(y)\, dy.$$

La fonction $g(y)$ étant arbitraire, de la dernière relation on déduit l'équation (50). ■

Remarque. Si un processus markovien $\xi(t)$ satisfait les conditions

$$\int_{\mathcal{R}^d} (y - x)\, \mathsf{P}(s, x, t, dy) = a(s, x)(t - s) + o(t - s), \tag{51}$$

$$\int_{\mathcal{R}^d} (z, y - x)^2\, \mathsf{P}(s, x, t, dy) = (b(s, x) z, z)(t - s) + o(t - s) \tag{52}$$

et pour un $\delta > 0$

$$\int_{\mathcal{R}^d} |y - x|^{2+\delta}\, \mathsf{P}(s, x, t, dy) = o(\Delta t), \tag{53}$$

alors il est de diffusion. En effet,

$$\mathsf{P}(s, x, t, \overline{S_\varepsilon(x)}) \leqslant \frac{1}{\varepsilon^{2+\delta}} \int_{\mathcal{R}^d} |y - x|^{2+\delta}\, \mathsf{P}(s, x, t, dy),$$

$$\int_{S_\varepsilon(x)} (y - x)\, \mathsf{P}(s, x, t, dy) = a(s, x)(t - s) + o(t - s) -$$
$$- \int_{\overline{S}_\varepsilon(x)} (y - x)\, \mathsf{P}(s, x, t, dy),$$

$$\int_{S_\varepsilon(x)} (z, y - x)^2\, \mathsf{P}(s, x, t, dy) = (b(s, x) z, z)(t - s) +$$
$$+ o(t - s) - \int_{\overline{S}_\varepsilon(x)} (z, y - x)^2\, \mathsf{P}(s, x, t, dy),$$

en outre pour $\alpha < 2+\delta$

$$\int\limits_{\bar{S}_\varepsilon(x)} |y-x|^{\alpha} \mathsf{P}(s, x, t, dy) \leqslant \frac{1}{\varepsilon^{2+\delta-\alpha}} \int\limits_{\bar{S}_\varepsilon(x)} |y-x|^{2+\delta} \mathsf{P}(s, x, t, dy).$$

Donc quel que soit $\varepsilon > 0$ fixe, les conditions (43), (44) et (45) sont réunies et le processus $\xi(t)$ est de diffusion.

§ 5. Processus stationnaires au sens large

Les processus stationnaires constituent une importante classe de processus aléatoires. On peut les définir comme des processus dont les caractéristiques probabilistes ne varient pas avec le temps ou encore comme des processus qui se déroulent dans des conditions ne variant pas avec le temps. Plus exactement, soit T un intervalle de temps fini ou non.

DÉFINITION. *Un processus aléatoire (au sens large) $\xi(t)$, $t \in T$, à valeurs dans $\mathcal{R}^d$ est stationnaire si quels que soient n et $t_1, t_2, \ldots, t_n$ tels que $t + t_k \in T$ $(k = 1, \ldots, n)$ la répartition conjointe des vecteurs aléatoires*

$$\xi(t_1 + t), \ldots, \xi(t_n + t) \tag{1}$$

ne dépend pas de t.

Il revient au même d'exiger que pour toute fonction continue bornée $f(x_1, \ldots, x_n)$, $x_k \in \mathcal{R}^d$, la quantité

$$\mathsf{E} f(\xi(t_1 + t), \ldots, \xi(t_n + t))$$

ne dépende pas de t. En particulier, si les coordonnées $\xi^j(t)$, $j = 1, \ldots, k$, du vecteur $\xi(t)$ admettent des moments finis du second ordre, les quantités

$$m^j(t) = \mathsf{E}\xi^j(t), \; j = 1, \ldots, d,$$

ne dépendent pas de t, $m^j(t) = m^j$, et

$$b^{jk}(t, s) = \mathsf{E}\xi^j(t)\,\xi^k(s), \; j, k = 1, \ldots, d, \; t \geqslant s,$$

ne dépendent que de la différence $t - s$: $b^{jk}(t, s) = b^{jk}(t - s)$.

Il existe un très grand nombre de problèmes relevant de la théorie des processus stationnaires dont la solution peut être exprimée en fonction des moments du premier et du second ordre des processus envisagés. C'est pourquoi il semble naturel de distinguer la classe de processus dont les moments du premier et du second ordre sont stationnaires. A. Khintchine [3] fut le premier à définir et à étudier cette classe.

DÉFINITION. *Un processus aléatoire $\xi(t) = (\xi^1(t), \ldots, \xi^d(t))$, $t \geqslant 0$, à valeurs dans $\mathcal{R}^d$ est stationnaire au sens large si $\mathsf{E}\,|\xi(t)|^2 < \infty$ et*

$$\mathsf{E}\xi(t) = m = \text{const}, \quad \mathsf{E}[\xi(t) - m]\,[\xi(s) - m]^* = R(t - s) \quad (t > s),$$

où $R(t)$ est une fonction matricielle continue.

La fonction $R(t)$ s'appelle *fonction de corrélation* (*matricielle*) du processus $\xi(t)$. Dans certains cas on a intérêt à étudier des processus aléatoires vectoriels à valeurs complexes $\zeta(t) = (\zeta^1(t), \ldots, \zeta^d(t))$, où $\zeta^k(t) = \xi^k(t) + i\eta^k(t)$, $\xi^k(t)$, $\eta^k(t)$ sont des processus aléatoires réels. Le processus $\zeta(t)$ est défini par la donnée d'un processus vectoriel $2d$-dimensionnel $\theta(t) = (\xi^1(t), \ldots, \xi^d(t), \eta^1(t), \ldots, \eta^d(t))$. Les répartitions de toutes les caractéristiques possibles de $\zeta(t)$ s'expriment alors sans peine en fonction des répartitions conjointes des vecteurs $\theta(t)$.

Comme exemple de processus stationnaire au sens large étudions des oscillations à paramètres aléatoires. On considérera des processus scalaires $\xi(t)$.

Introduisons quelques définitions qui nous seront utiles pour l'interprétation physique des processus aléatoires. Si $\xi(t)$ est l'intensité d'un courant à l'instant t et que l'on s'intéresse à l'énergie dissipée par une résistance unitaire, les définitions suivantes sont naturelles.

On appelle *énergie* transférée par le processus aléatoire $\xi(t)$ pendant l'intervalle de temps $]t_1, t_2[$ la quantité

$$\int_{t_1}^{t_2} \xi^2(t)\,dt;$$

énergie totale transférée par le processus $\xi(t)$ $(-\infty < t < \infty)$, l'intégrale $\int_{-\infty}^{\infty} \xi^2(t)\,dt$ si elle existe.

On appelle *puissance moyenne* d'un processus aléatoire la limite

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \xi^2(t)\,dt.$$

Si le processus $\xi(t)$ est à valeurs complexes, il faudra remplacer $\xi^2(t)$ par $|\xi(t)|^2$ dans les expressions précédentes.

Dans la suite nous adopterons cette terminologie même si le processus auquel on l'applique admet une autre interprétation physique.

Dans de nombreux problèmes les processus aléatoires sont simulés par des sommes d'harmoniques à fréquences données et à amplitudes et phases aléatoires. En d'autres termes, on considère des processus de la forme

$$\xi(t) = \sum_{k=1}^{n} \alpha_k \cos(u_k t + \varphi_k),$$

où u_k sont des nombres donnés, α_k et φ_k des variables aléatoires. La structure probabiliste de ce processus est complètement définie par la répartition conjointe des variables aléatoires α_k, φ_k ($k = 1, 2, \ldots, n$). Ceci étant, par processus défini par la formule (1) on entend un processus aléatoire dont les répartitions finidimensionnelles sont susceptibles d'être calculées à l'aide de la formule (1) et de la répartition conjointe des variables $\alpha_1, \ldots, \alpha_n, \varphi_1, \ldots, \varphi_n$.

Le processus aléatoire $\xi(t)$ est la somme de n oscillations harmoniques d'amplitudes $|\alpha_k|$ et de fréquences u_k.

Dans de nombreux cas il y a intérêt à considérer des processus aléatoires à valeurs complexes à caractère oscillatoire

$$\zeta(t) = \sum_{k=1}^{n} \gamma_k e^{iu_k t}, \tag{2}$$

où les amplitudes complexes γ_k sont aléatoires:

$$\gamma_k = \alpha_k + i\beta_k,$$

α_k, β_k, $k = 1, 2, \ldots, n$, sont des réels. L'ensemble de toutes les fréquences $\{u_k\}$, $k = 1, 2, \ldots, n$, considérées comme des points de la droite $(-\infty < u < \infty)$, s'appelle *spectre* de la fonction aléatoire $\zeta(t)$.

Le processus $\zeta(t)$ peut être décomposé en une partie réelle et une partie imaginaire:

$$\zeta(t) = \xi(t) + i\eta(t),$$

où

$$\begin{aligned} \xi(t) &= \sum_{k=1}^{n} \alpha_k \cos u_k t - \beta_k \sin u_k t, \\ \eta(t) &= \sum_{k=1}^{n} \alpha_k \sin u_k t + \beta_k \cos u_k t. \end{aligned} \tag{3}$$

On calcule sans peine la puissance moyenne transférée par le processus $\zeta(t)$. On a

$$\frac{1}{2T}\int_{-T}^{T} |\zeta(t)|^2\,dt = \frac{1}{2T}\int_{-T}^{T} \sum_{k,\,r=1}^{n} \gamma_k \bar{\gamma}_r e^{it(u_k-u_r)}\,dt =$$

$$= \sum_{k=1}^{n} |\gamma_k|^2 + \sum_{\substack{k,\,r=1 \\ k\neq r}}^{n} \gamma_k \bar{\gamma}_r \frac{\sin T(u_k-u_r)}{T(u_k-u_r)}.$$

Lorsque $T \to \infty$, on obtient

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} |\zeta(t)|^2\,dt = \sum_{k=1}^{n} |\gamma_k|^2.$$

Donc la puissance moyenne transférée par le processus aléatoire oscillatoire $\zeta(t)$ est égale à la somme des puissances moyennes transférées par chaque composante harmonique du processus.

On calcule de façon analogue la valeur moyenne de la fonction aléatoire $\zeta(t)$ sur un intervalle de temps infini. Il vient

$$\frac{1}{2T}\int_{-T}^{T}\zeta(t)\,dt=\sum_{k=1}^{n}\gamma_k\frac{\sin Tu_k}{Tu_k};$$

si le point 0 est un point du spectre du processus aléatoire, la fonction $\frac{\sin u}{u}$ est supposée égale à 1 en ce point ($u=0$). Donc

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\zeta(t)\,dt=\gamma_0,$$

où γ_0 est l'amplitude correspondant à la fréquence $u=0$.

La répartition de la variable $\zeta(t)$ est très compliquée même sous des hypothèses spéciales sur la répartition des variables γ_k. Cependant les caractéristiques élémentaires de la répartition de la variable aléatoire $\zeta(t)$ ne sont pas difficiles à établir. Supposons que les amplitudes complexes γ_k possèdent des espérances mathématiques nulles et qu'elles ne sont pas corrélées, c'est-à-dire

$$\mathsf{E}\gamma_k=0,\quad \mathsf{E}\gamma_k\overline{\gamma_r}=0,\quad k\neq r.$$

On a alors

$$\mathsf{E}\zeta(t)=0.$$

On remarquera que si 0 n'est pas point du spectre du processus aléatoire, l'espérance mathématique de la fonction $\zeta(t)$, c'est-à-dire sa valeur moyenne au sens probabiliste, coïncide avec sa valeur moyenne prise sur l'intervalle de temps infini $]-\infty, \infty[$. Si au contraire 0 est point du spectre, la moyenne temporelle de la fonction échantillonnée est une variable aléatoire. On a l'expression suivante pour la fonction de corrélation du processus $\zeta(t)$:

$$R(t_1,t_2)=\mathsf{E}[\zeta(t_1)\overline{\zeta(t_2)}]=\mathsf{E}\Big[\sum_{k,\,r=1}^{n}\gamma_k\overline{\gamma_r}\,e^{i(u_kt_1-u_rt_2)}\Big]=\sum_{k=1}^{n}c_k^2e^{iu_k(t_1-t_2)},$$

où $c_k^2=\mathsf{E}|\gamma_k|^2$. Par suite, la fonction de corrélation du processus $\zeta(t)$ dépend uniquement de la différence t_1-t_2:

$$R(t_1,t_2)=R(t_1-t_2),\tag{4}$$

$$R(t)=\sum_{k=1}^{n}c_k^2e^{iu_kt}.\tag{5}$$

Si donc les variables γ_k du processus (2) ne sont pas corrélées et possèdent des valeurs moyennes nulles, ce processus est stationnaire

au sens large et sa fonction de corrélation est donnée par (5). Cette formule s'appelle *représentation spectrale* de la fonction de corrélation. Elle définit le spectre du processus aléatoire, c'est-à-dire l'ensemble des fréquences $\{u_k\}$, $k = 1, \ldots, n$, des oscillations harmoniques constituant le processus $\zeta(t)$ et des espérances mathématiques c_k^2 des puissances moyennes transférées par les composantes respectives du processus. La quantité c_k^2 s'appelle *valeur moyenne de la puissance* de la composante harmonique du processus de fréquence u_k. Elle s'obtient par centrage temporel de la puissance et ensuite centrage au sens probabiliste. Ces considérations énergétiques suggèrent l'introduction de la fonction spectrale qui est une caractéristique importante du processus stationnaire.

La *fonction spectrale* $F(u)$ du processus (2) est définie par

$$F(u) = \sum_{\substack{k \\ u_k < u}} c_k^2.$$

Ceci signifie que $F(u)$ est égale à la puissance moyenne transférée par les composantes harmoniques du processus $\zeta(t)$ dont les fréquences sont inférieures à la valeur donnée u. Elle caractérise complètement tant la puissance moyenne de chaque harmonique de $\zeta(t)$ que la puissance moyenne globale des harmoniques dont les fréquences sont contenues dans un intervalle quelconque donné. En effet,

$$c_k^2 = F(u_k + 0) - F(u_k), \qquad \sum_{u_1 \leqslant u_k \leqslant u_2} c_k^2 = F(u_2) - F(u_1).$$

En utilisant la fonction spectrale on peut mettre la fonction de corrélation de $\zeta(t)$ sous la forme

$$R(t) = \int_{-\infty}^{\infty} e^{iut}\, dF(u). \tag{6}$$

D'un point de vue mathématique, la fonction spectrale est une fonction continue à gauche, non croissante, non négative, partout constante, sauf en un nombre fini de points où elle présente des sauts de c_k^2. Il se trouve que la notion de fonction spectrale peut être étendue aux processus stationnaires au sens large. Ce problème ainsi que la généralisation de (6) à des processus stationnaires arbitraires seront étudiés plus bas.

Il semble particulièrement intéressant d'étudier les processus de la forme (2) par passage à la limite. Ce passage consiste à faire croître indéfiniment le nombre de termes de la somme (2) lorsque les amplitudes complexes γ_k décroissent; le spectre du processus, c'est-à-dire l'ensemble de toutes les fréquences u_k, devient à la limite partout dense dans la droite $]-\infty, \infty[$. On obtient un processus aléatoire dont on dit qu'il possède un spectre continu. Plus

bas et au chapitre V on donnera une définition plus exacte de ce terme et l'on examinera l'analogue de la représentation (2) pour les processus ainsi obtenus. Pour l'instant on se limitera à l'étude des répartitions limites.

THÉORÈME 1. *Soient* α_{nk}, β_{nk} $(k = 1, \ldots, n;\ n = 1, 2, \ldots)$ *deux suites de séries de variables aléatoires deux à deux indépendantes (dans chaque série)*

$$\mathsf{E}\alpha_{nk} = \mathsf{E}\beta_{nk} = 0, \quad \operatorname{Var}\alpha_{nk} = \operatorname{Var}\beta_{nk} = \frac{b_{nk}^2}{2} < \infty$$

et

$$\zeta_n(t) = \sum_{k=1}^{n} \gamma_{nk} e^{iu_{nk}t} = \xi_n(t) + i\eta_n(t), \quad \gamma_{nk} = \alpha_{nk} + i\beta_{nk}.$$

Supposons que, lorsque $n \to \infty$, *les conditions suivantes sont réunies:*

a) *les fonctions spectrales* $F_n(u)$ *des processus* $\zeta_n(t)$ *convergent vers une fonction* $F(u)$ *sur un ensemble de valeurs partout dense dans la droite* $]-\infty, \infty[$ *et que*

$$F(-\infty) = 0, \quad F(+\infty) = \lim \sum_{k=1}^{n} b_{nk}^2 < \infty;$$

b) *les variables aléatoires* $\{\alpha_{nk}, \beta_{nk}\}$ *satisfont la condition de Lindeberg* (théorème 5, § 2).

Sous ces conditions le processus aléatoire $\zeta_n(t)$ *converge faiblement vers un processus stationnaire* $\zeta(t) = \xi(t) + i\eta(t)$ *lorsque* $n \to \infty$. *La fonction caractéristique de la répartition conjointe des variables*

$$\xi(t_1), \ldots, \xi(t_s), \eta(t_1), \ldots, \eta(t_s) \tag{7}$$

s'écrit

$$\varphi(u_1, \ldots, u_s, v_1, \ldots, v_s) = \exp\left\{-\frac{1}{2}B^2\right\},$$

où

$$B^2 = \frac{1}{2} \sum_{j,\,k=1}^{s} R(t_j - t_k) z_j \bar{z}_k, \quad z_j = u_j - iv_j, \quad j = 1, \ldots, s,$$

$$R(t) = \int_{-\infty}^{\infty} e^{iut}\, dF(u). \tag{8}$$

On remarquera que par hypothèse $F(u)$ est une fonction monotone non croissante bornée.

Le théorème formulé est un cas particulier du théorème 5, § 2. On admettra que Θ est l'ensemble des couples $\theta = (t, q)$, $-\infty < t < \infty$, $q = 1, 2$, $\alpha_{nk}(\theta) = \alpha_{nk}(t, q)$, où

$$\alpha_{nk}(t, 1) = \operatorname{Re} \gamma_{nk} e^{iu_{nk}t}, \quad \alpha_{nk}(t, 2) = \operatorname{Im} \gamma_{nk} e^{iu_{nk}t}.$$

Il est aisé de vérifier que si la condition de Lindeberg est réalisée par α_{nk}, β_{nk}, elle le sera par $\alpha_{nk}(\theta)$ quel que soit θ. Soient

$$R^{(n)}(\theta_1, \theta_2) = R^{(n)}_{q_1 q_2}(t_1, t_2), \quad \theta_i = (t_i, q_i),$$
$$R^{(n)}_{11}(t_1, t_2) = \mathsf{E}\xi_n(t_1)\,\xi_n(t_2), \quad R^{(n)}_{22}(t_1, t_2) = \mathsf{E}\eta_n(t_1)\,\eta_n(t_2),$$
$$R^{(n)}_{12}(t_1, t_2) = \mathsf{E}\xi_n(t_1)\,\eta_n(t_2).$$

La formule (3) entraîne

$$R^{(n)}_{11}(t_1, t_2) = R^{(n)}_{22}(t_1, t_2) = \frac{1}{2}\int_{-\infty}^{\infty} \cos[(t_2 - t_1)u]\, dF_n(u),$$

$$R^{(n)}_{12}(t_1, t_2) = \frac{1}{2}\int_{-\infty}^{\infty} \sin[(t_2 - t_1)u]\, dF_n(u).$$

Le théorème de Helly et la condition a) entraînent l'existence des limites

$$R_{ij}(t) = \lim_{n\to\infty} R^{(n)}_{ij}(\tau, \tau + t), \quad i, j = 1, 2,$$

$$R_{11}(t) = R_{22}(t) = \frac{1}{2}\int_{-\infty}^{\infty} \cos(tu)\, dF(u),$$

$$R_{12}(t) = \frac{1}{2}\int_{-\infty}^{\infty} \sin(tu)\, dF(u).$$

Donc les hypothèses du théorème 5, § 2, sont réunies. Ce théorème nous dit que la fonction caractéristique de la répartition conjointe des variables (7) vaut

$$\varphi(u_1, \ldots, u_s, v_1, \ldots, v_s) = e^{-\frac{1}{2}B^2},$$

où

$$B^2 = \sum_{k, j=1}^{s} R_{11}(t_j - t_k)\, u_k u_j + 2R_{12}(t_j - t_k)\, u_k v_j + R_{22}(t_j - t_k)\, v_k v_j =$$
$$= \frac{1}{2}\sum_{j, k=1}^{s} R(t_j - t_k)\, z_j \bar{z}_k,$$

$$z_j = u_j - iv_j, \quad j = 1, \ldots, s. \blacksquare$$

Le théorème démontré nous place dans le cas d'un processus stationnaire dont la fonction de corrélation est donnée par la formule (8), où $F(u)$ est une fonction monotone, non croissante, bornée, non négative, continue à gauche.

On constate que l'expression (8) est générique pour tous les processus stationnaires au sens large. On a précisément le

THÉORÈME 2 (théorème de Khintchine). *Pour qu'une fonction continue $R(t)$ $(-\infty < t < \infty)$ soit fonction de corrélation d'un processus stationnaire au sens large il est nécessaire et suffisant qu'elle admette la représentation* (8).

Démonstration. *Condition nécessaire.* La fonction de corrélation d'un processus stationnaire au sens large est définie positive, c'est-à-dire

$$\sum_{k,\,j=1}^{n} R(t_k - t_j)\,\lambda_k \bar{\lambda}_j \geqslant 0$$

quels que soient n, $t_1, \ldots, t_n$, $\lambda_1, \ldots, \lambda_n$ (n est entier, t_k réels, λ_k complexes). Le théorème de Bochner-Khintchine entraîne que $R(t)$ admet la représentation (8). La *condition suffisante* découle du théorème précédent. ■

Du reste on peut considérablement simplifier la construction d'un processus stationnaire au sens large dont la fonction de corrélation est exprimée par la formule (8) avec une fonction spectrale $F(u)$ donnée *a priori*.

Supposons que $F(+\infty) = \sigma^2$ et que ξ est une variable aléatoire de fonction de répartition $\frac{1}{\sigma^2} F(x)$. Posons

$$\zeta(t) = \sigma e^{i(t\xi + \varphi)},$$

où φ et ξ sont indépendantes, φ étant uniformément répartie sur l'intervalle $]\,0,\ 2\pi\,[$. Il vient alors

$$\mathsf{E}\zeta(t) = \sigma \int_{-\infty}^{\infty} \int_{0}^{2\pi} e^{i(tu+\varphi)} \frac{1}{\sigma^2}\, dF(u)\, \frac{d\varphi}{2\pi} = 0$$

et

$$R(t+h, h) = \mathsf{E}\{\zeta(t+h)\,\overline{\zeta(h)}\} =$$

$$= \sigma^2 \int_{-\infty}^{\infty} \int_{0}^{2\pi} e^{itu} \frac{1}{\sigma^2}\, dF(u)\, \frac{d\varphi}{2\pi} = \int_{-\infty}^{\infty} e^{itu}\, dF(u).$$

DÉFINITION. *La fonction $F(u)$ qui figure dans la représentation* (8) *de la fonction de corrélation du processus stationnaire (au sens large) s'appelle fonction spectrale. Si $F(u)$ est absolument continue,*

$$F(u) = \int_{-\infty}^{u} f(u)\, du,$$

$f(u)$ est dite densité spectrale du processus.

Rappelons la signification physique de la fonction spectrale. Si la fonction aléatoire $\xi(t)$ est assimilée à un courant électrique, la fonction $F(u)$ peut être interprétée de la façon suivante : représentons le processus $\xi(t)$ sous forme d'une « somme indéfinie » d'oscillations harmoniques simples à amplitudes aléatoires et posons l'accroissement $F(u_2) - F(u_1)$ $(u_1 < u_2)$ égal à la puissance moyenne dissipée par les composantes harmoniques du processus dont les fréquences sont comprises dans l'intervalle $[u_1, u_2[$.

A noter que la fonction $R^{-1}(0) R(t)$, où $R(0) = F(+\infty)$, peut être considérée comme la fonction caractéristique de la fonction de répartition $(F(+\infty))^{-1} F(x)$. D'où il suit que la fonction spectrale est définie de façon unique par la fonction de corrélation. Si a et b sont des points de continuité de la fonction de répartition, on sait de la théorie des fonctions caractéristiques que

$$F(b) - F(a) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-ita} - e^{-itb}}{it} R(t)\, dt, \tag{9}$$

où l'intégrale est prise dans sa valeur principale. Aux points de discontinuité de la fonction $F(u)$ la formule (9) reste valable sous réserve que l'on remplace $F(u)$ par $\frac{F(u+0)+F(u)}{2}$.

Si la fonction de corrélation $R(t)$ est absolument intégrable sur $]-\infty, \infty[$, c'est-à-dire $\int_{-\infty}^{\infty} |R(t)|\, dt < \infty$, le second membre de (9) est dérivable par rapport au paramètre b. Donc l'intégrabilité absolue de la fonction $R(t)$ entraîne l'existence de la densité spectrale et

$$f(u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iut} R(t)\, dt,$$

c'est-à-dire $f(u)$ est l'inverse de $R(t)$ par la transformation de Fourier.

Le théorème 2 admet plusieurs généralisations. D'abord on peut considérer les processus stationnaires vectoriels (au sens large), ensuite les fonctions stationnaires (scalaires ou vectorielles) de plusieurs variables. Voyons dans un premier temps les processus stationnaires vectoriels au sens large à composantes complexes. Soit $\zeta(t) = (\zeta_1(t), \ldots, \zeta_d(t))$.

THÉORÈME 3. *Une fonction matricielle continue $R(t) = (R_{jk}(t))$, $j, k = 1, \ldots, d$, est fonction matricielle de corrélation d'un processus stationnaire au sens large $\zeta(t)$ si et seulement si on peut la mettre sous*

la forme

$$R(t) = \int_{-\infty}^{\infty} e^{iut}\, dF(u), \tag{10}$$

où $F(u) = (F_{jk}(u))$, $j, k, = 1, \ldots, d$, *est une fonction matricielle telle que*

a) *quels que soient* u_1 *et* u_2, $u_1 < u_2$, *la matrice* $\Delta F(u) = F(u_2) - F(u_1)$ *est définie non négative*;

b) $\mathrm{Tr}\,\{F(+\infty) - F(-\infty)\} < \infty$.

S'agissant de la condition b), on remarquera qu'en vertu de a) les éléments diagonaux de la matrice $F(u)$ sont des fonctions de u monotones non décroissantes. Donc il revient au même d'exiger que chaque élément diagonal $F_{ii}(u)$ de la matrice $F(u)$ soit à variation bornée sur la droite $]-\infty, \infty[$.

D'autre part, la définition positive de la matrice $\Delta F(u)$ entraîne

$$|\Delta F_{jk}(u)|^2 \leqslant \Delta F_{jj}(u)\, \Delta F_{kk}(u),$$

d'où

$$\sum_{p=1}^{s} |\Delta F_{jk}(u_p)| \leqslant \left[\sum_{p=1}^{s} \Delta F_{jj}(u_p)\right]^{\frac{1}{2}} \left[\sum_{p=1}^{s} \Delta F_{kk}(u_p)\right]^{\frac{1}{2}},$$

où $\Delta F(u_p) = F(u_p) - F(u_{p-1})$, $p = 1, 2, \ldots, s$, $u_0 < u_1 < \ldots < u_s$.

De là il suit que les éléments non diagonaux $F_{jk}(u)$ de $F(u)$ seront des fonctions à variation bornée. Sans restreindre la généralité on peut admettre que $F(-\infty) = 0$.

Le théorème 3 découle de la variante matricielle du théorème de Bochner-Khintchine. On peut le déduire comme corollaire du théorème de Bochner-Khintchine. Prouvons ce théorème pour illustrer le passage des théorèmes « unidimensionnels » à leurs analogues « multidimensionnels ».

D é m o n s t r a t i o n d u t h é o r è m e 3. Soit $R(t)$ la fonction matricielle de corrélation continue d'un processus vectoriel stationnaire au sens large d-dimensionnel à composantes complexes. Pour tout vecteur complexe d-dimensionnel on considère la variable aléatoire

$$\zeta_c(t) = (\zeta(t), c).$$

$\zeta_c(t)$ est alors un processus aléatoire stationnaire au sens large à fonction de corrélation continue:

$$m_c = \mathsf{E}\zeta_c(t) = (\mathsf{E}\zeta(t), c) = \text{const},$$

$$R_c(t)\, \mathsf{E} \underset{\text{Déf}}{=} \mathsf{E}([\zeta_c(t+s) - m_c]\, \overline{[\zeta_c(s) - m_c]}) = c^* R(t)\, c, \quad t > 0.$$

En vertu du théorème 2, la fonction de corrélation $R_c(t)$ peut être mise sous la forme

$$R_c(t)=\int_{-\infty}^{\infty} e^{iut}\,dF_c(u),$$

où $F_c(u)$ est une fonction monotone non croissante de u et $F_c(u) < < \infty$.

Soit $e^{(k)}$ un vecteur d-dimensionnel tel que

$$e_p^{(k)}=\begin{cases}0, & p\neq k,\\ 1, & p=k.\end{cases}$$

$R_{e^{(k)}}(t) = R_{kk}(t)$ est alors la fonction de corrélation de la k-ième composante du processus vectoriel $\zeta(t)$. Posons

$$e^{(k,\,j)}=e^{(k)}+e^{(j)},\quad \tilde{e}^{(k,\,j)}=ie^{(k)}+e^{(j)}.$$

Il vient

$$R_{e^{(k,\,j)}}(t)=R_{kk}(t)+R_{kj}(t)+R_{jk}(t)+R_{jj}(t),\quad k\neq j,$$

$$R_{\tilde{e}^{(k,\,j)}}(t)=R_{kk}(t)+iR_{kj}(t)-iR_{jk}(t)+R_{jj}(t),$$

d'où

$$R_{kj}(t)=\frac{R_{e^{(k,\,j)}}(t)-iR_{\tilde{e}^{(k,\,j)}}(t)}{2}-\frac{1-i}{2}\left(R_{e^{(k)}}-R_{e^{(j)}}\right).$$

Posons

$$F_{kj}(u)=\frac{F_{e^{(k,\,j)}}(u)-iF_{\tilde{e}^{(k,\,j)}}(u)}{2}-\frac{1-i}{2}\left(F_{e^{(k)}}(u)-F_{e^{(j)}}(u)\right),\quad k\neq j,$$

$$F_{kk}(u)=F_{e^{(k)}}(u),$$

on a

$$R_{kj}(t)=\int_{-\infty}^{\infty} e^{itu}\,dF_{kj}(u),\quad k,\ j=1,\ \ldots,\ d.$$

Il vient

$$R_c(t)=\sum_{k,\,j=1}^{d}\bar{c}_kR_{kj}c_j=\int_{-\infty}^{\infty} e^{itu}d\left(\sum_{k,\,j=1}^{d}\bar{c}_kF_{kj}(t)\,c_j\right).$$

L'unicité de la fonction spectrale $F_c(u)$ entraîne

$$F_c(u)=\sum_{k,\,j=1}^{d}\bar{c}_kF_{kj}(u)\,c_j.$$

D'où il suit que $\Delta F_c(u) = \sum_{k,\, j=1}^{d} \bar{c}_k \Delta F_{kj}(u) c_j \geqslant 0$, de sorte que la matrice $\Delta F(u)$ est définie non négative et $F_{kk}(+\infty) < \infty$. Reste maintenant à démontrer qu'à toute fonction matricielle $F(u)$ possédant les propriétés a) et b) du théorème 3 on peut associer un processus vectoriel stationnaire au sens large de fonction de corrélation (10). A cet effet introduisons tout d'abord un processus gaussien réel $2d$-dimensionnel $(\xi(t), \eta(t))$, où $\xi(t)$ et $\eta(t)$ sont des processus gaussiens réels d-dimensionnels, que l'on définira comme suit. Soient $\bar{a}_k = (a_k^1, \ldots, a_k^d)$, $\bar{b}_k = (b_k^1, \ldots, b_k^d)$, $k = 1, \ldots, n$, deux suites de vecteurs dans $\mathcal{R}^d$. Définissons la fonction caractéristique de la répartition conjointe des vecteurs $(\xi(t_1), \ldots, \xi(t_n), \eta(t_1), \ldots, \eta(t_n))$ avec la formule

$$\varphi_{t_1, \ldots, t_n}(\bar{a}_1, \ldots, \bar{a}_n, \bar{b}_1, \ldots, \bar{b}_n) = \\ = \exp\left\{-\frac{1}{2} K(\bar{a}_1, \ldots, \bar{a}_n, \bar{b}_1, \ldots, \bar{b}_n\right\},$$

où

$$K(\bar{a}_1, \ldots, \bar{a}_n, \bar{b}_1, \ldots, \bar{b}_n) = \sum_{k,\, j=1}^{n} \bar{a}_k^* R'(t_k - t_j) \bar{a}_j + \\ + \sum_{k,\, j=1}^{n} [a_k^* R''(t_k - t_j) \bar{b}_j + \bar{b}_k^* R''(t_k - t_j) \bar{a}_j] + \sum_{k,\, j=1}^{n} \bar{b}_k^* R'(t_k - t_j) \bar{b}_j,$$

$$R'(t) = \frac{1}{2} \int_{-\infty}^{\infty} \cos(tu)\, dF(u), \quad R''(t) = -\frac{1}{2} \int_{-\infty}^{\infty} \sin(tu)\, dF(u),$$

$$R'(t) - iR''(t) = \frac{1}{2} R(t) = \frac{1}{2} \int_{-\infty}^{\infty} e^{itu}\, dF(u).$$

Pour que cette définition ait un sens, c'est-à-dire pour que la fonction $\varphi_{t_1, \ldots, t_n}(\bar{a}_1, \ldots, \bar{a}_n, \bar{b}_1, \ldots, \bar{b}_n)$ soit effectivement fonction caractéristique de la répartition conjointe de $2dn$ variables aléatoires, il est nécessaire et suffisant que la fonction $K(\bar{a}_1, \ldots, \bar{a}_n, \bar{b}_1, \ldots, \bar{b}_n)$ soit une forme quadratique définie non négative. On voit sans peine que

$$K(\bar{a}_1, \ldots, \bar{a}_n, \bar{b}_1, \ldots, \bar{b}_n) = \frac{1}{2} \operatorname{Re} \sum z_k^* R(t_k - t_j) z_j,$$

où $z_k = a_k - ib_k$, $k = 1, \ldots, n$. Par ailleurs

$$\sum z_k^* R(t_k - t_j) z_j = \int_{-\infty}^{\infty} \left(\sum_{k=1}^{n} e^{-it_k u} z_k\right)^* dF(u) \left(\sum_{k=1}^{n} e^{-it_k u} z_k\right) \geqslant 0.$$

Donc $\varphi_{t_1, \ldots, t_n}$ est bien fonction caractéristique. On s'assure sans peine que la famille de répartitions $\{F_{t_1, \ldots, t_n}\}$ associées aux fonctions caractéristiques $\varphi_{t_1, \ldots, t_n}$ vérifie les conditions de compatibilité (§ 1), de sorte qu'elle définit un processus aléatoire gaussien au sens large. Ceci étant,

$$\mathsf{E}\xi(t) = \mathsf{E}\eta(t) = 0, \quad \mathsf{E}\xi(t)\xi(s) = \mathsf{E}\eta(t)\eta(s) = R'(t, s),$$
$$\mathsf{E}\xi(t)\eta(s) = R''(t-s).$$

Posons $\zeta(t) = \xi(t) + i\eta(t)$. Le processus $\zeta(t)$ est un processus gaussien vectoriel à composantes complexes. On a

$$\mathsf{E}\zeta(t) = 0, \quad \mathsf{E}\zeta(t)\zeta^*(s) = \mathsf{E}[\xi(t)\xi^*(s) + \eta(t)\eta^*(s)] -$$
$$-i\mathsf{E}[\xi(t)\eta^*(s) + \eta(t)\xi^*(s)] = 2R(t-s) - 2iR''(t-s) = R(t-s).$$

Donc $\xi(t)$ est un processus gaussien stationnaire au sens large à fonction matricielle de corrélation $R(t-s)$. ■

La notion de stationnarité au sens large peut être généralisée aux fonctions aléatoires de plusieurs variables. Soit $\zeta(x) = (\zeta^1(x), \ldots, \zeta^d(x))$ une fonction aléatoire vectorielle au sens large, éventuellement à coordonnées complexes, définie pour tous les $x \in \mathcal{R}^m$. On l'appellera *champ aléatoire*. Un champ aléatoire $\zeta(x)$ est *homogène* si

$$\mathsf{E}\zeta(x) = m = \text{const}, \quad \mathsf{E}(\zeta(x) - m)(\zeta(y) - m)^* = R(x-y),$$

où $R(x)$ est une fonction matricielle continue appelée *fonction de corrélation du champ homogène*. La fonction matricielle $R(x)$ est définie non négative. Ce qui veut dire que quels que soient les vecteurs z_k complexes d-dimensionnels, les points $x_k \in \mathcal{R}^m$, $k = 1, \ldots, n$ et n entier quelconque, on a

$$\sum_{k, j=1}^{n} z_k^* R(x_k - x_j) z_j \geqslant 0.$$

Le théorème de Bochner-Khintchine se généralise facilement aux fonctions définies non négatives de $x \in \mathcal{R}^m$. Par analogie au théorème 3 on démontre le

THÉORÈME 4. *Pour qu'une fonction matricielle $R(x)$, $x \in \mathcal{R}^d$,* soit *fonction de corrélation d'un champ homogène, il est nécessaire et suffisant qu'elle admette la représentation*

$$R(x) = \int_{\mathcal{R}^d} e^{i(x, u)} F(du), \tag{11}$$

où $F(A)$ est une fonction d'ensembles matricielle complexe dénombrablement additive définie sur les boréliens de $\mathcal{R}^m$ et telle que $z^*F(A)z \geqslant 0$ quels que soient z, vecteur complexe, et A, ensemble de $\mathfrak{B}^m$. En outre $\operatorname{Tr} F(\mathcal{R}^m) < \infty$.

Un champ aléatoire $\zeta(x)$ est dit *homogène* et *isotrope* si sa fonction de corrélation $R(x)$ dépend uniquement de la longueur du vecteur x. Donc pour un champ homogène et isotrope

$$\mathsf{E}(\zeta(x)-m)(\zeta(y)-m)^* = R(\rho),$$

où $\rho = \rho(x, y)$ est la distance entre les points x et y.

Trouvons l'expression de la fonction de corrélation d'un champ homogène et isotrope. A cet effet considérons l'expression de la fonction de corrélation d'un champ homogène

$$R(\rho) = \int\limits_{\mathcal{R}^m} e^{i(x,u)} F(du)$$

et étendons l'intégration à la sphère S_ρ de rayon ρ. En intervertissant l'ordre d'intégration on obtient

$$R(\rho) = \frac{\Gamma\left(\frac{m}{2}\right)}{2\pi^{\frac{m}{2}}\rho^{m-1}} \int\limits_{\mathcal{R}^m} \left\{ \int\limits_{S_\rho} e^{i(x,u)}\, ds \right\} F(du). \tag{12}$$

Soient $f(x)$ fonction intégrable sur $\mathcal{R}^m$, V_ρ boule de rayon ρ centrée en un point fixe. Il vient

$$\frac{d}{d\rho} \int\limits_{V_\rho} f(x)\, dx^1 \ldots dx^m = \int\limits_{S_\rho} f(x)\, ds,$$

où au second membre l'intégration est étendue à la surface S_ρ de la boule V_ρ. Appliquons cette formule au calcul de l'intégrale dans l'accolade du second membre de (12). En passant aux coordonnées sphériques dans un espace n-dimensionnel et en désignant par φ_1 l'angle des vecteurs x et u on obtient

$$\int\limits_{V_\rho} e^{i(x,u)}\, dx^1 \ldots dx^m = \int\limits_0^\rho \int\limits_0^\pi \ldots \int\limits_0^\pi \int\limits_0^{2\pi} e^{ir|u|\cos\varphi_1} r^{m-1} \sin^{m-2}\varphi_1 \times$$

$$\times \sin^{m-3}\varphi_2 \ldots \sin\varphi_{m-2}\, dr\, d\varphi_1 \ldots d\varphi_{m-2} =$$

$$= \frac{2\pi^{\frac{m-1}{2}}}{\Gamma\left(\frac{m-1}{2}\right)} \int\limits_0^\rho \int\limits_0^\pi e^{ir|u|\cos\varphi_1} r^{m-1} \sin^{m-2}\varphi_1\, dr\, d\varphi_1.$$

D'autre part

$$\int\limits_0^\pi e^{ir|u|\cos\varphi_1} \sin^{m-2}\varphi_1\, d\varphi_1 = \sum_{k=0}^\infty \int\limits_0^\pi \frac{(ir|u|\cos\varphi_1)^k}{k!} \sin^{m-2}\varphi_1\, d\varphi_1 =$$

$$= \sum_{k=0}^\infty (-1)^k \frac{r^{2k}|u|^{2k}}{(2k)!} \frac{\Gamma\left(\frac{2k+1}{2}\right)\Gamma\left(\frac{m-1}{2}\right)}{\Gamma\left(\frac{2k+m}{2}\right)}$$

et

$$I=\int_0^\rho\int_0^\pi e^{ir\,|u|\cos\varphi_1}r^{m-1}\sin^{m-2}\varphi_1\,d\varphi_1=$$

$$=\sum_{k=0}^{\infty}(-1)^k\frac{\rho^{2k+m}\,|u|^{2k}}{(2k)!\,(2k+m)}\,\frac{\Gamma\left(\frac{2k+1}{2}\right)\Gamma\left(\frac{m-1}{2}\right)}{\Gamma\left(\frac{2k+m}{2}\right)}.$$

En vertu de la formule

$$\Gamma\left(k+\frac{1}{2}\right)=\frac{\sqrt{\pi}}{2^{2k-1}}\,\frac{\Gamma(2k)}{\Gamma(k)},$$

on a

$$I=\frac{\sqrt{\pi}\,\Gamma\left(\frac{m-1}{2}\right)2^{\frac{m}{2}-1}\rho^{\frac{m}{2}}}{|u|^{\frac{m}{2}}}\sum_{k=0}^{\infty}(-1)^k\frac{\left(\frac{\rho|u|}{2}\right)^{2k+\frac{m}{2}}}{k!\,\Gamma\left(k+\frac{m}{2}+1\right)}=$$

$$=\frac{\sqrt{\pi}}{2}\Gamma\left(\frac{m-1}{2}\right)\left(\frac{2\rho}{|u|}\right)^{\frac{m}{2}}J_{\frac{m}{2}}(\rho|u|),$$

où $J_n(x)$ est une fonction de Bessel du premier genre d'ordre n. Donc

$$\int_{V_\rho}e^{i(x,\,u)}\,dx^1\,\ldots\,dx^n=\left(\frac{2\pi\rho}{|u|}\right)^{\frac{m}{2}}J_{\frac{m}{2}}(\rho|u|).$$

D'où il suit

$$\int_{S_\rho}e^{i(x,\,u)}\,ds=\left(\frac{2\pi\rho}{|u|}\right)^{\frac{m}{2}}|u|\,J_{\frac{m-2}{2}}(\rho|u|).$$

En particulier, l'intégrale envisagée dépend de $|u|$. Introduisons le paramètre positif λ et posons

$$g(\lambda)=F(V_\lambda),\quad\lambda>0.$$

La dernière formule et la formule (12) donnent

$$R(\rho)=2^{\frac{m-2}{2}}\Gamma\left(\frac{m}{2}\right)\int_0^\infty\frac{J_{\frac{m-2}{2}}(\lambda\rho)}{(\lambda\rho)^{\frac{m-2}{2}}}\,dg(\lambda),\tag{13}$$

où $g(\lambda)$ est une fonction monotone croissante, $g(-0)=0$ et

$$g(+\infty)=F(\mathscr{R}^m)=R(0)<\infty.$$

Nous avons donc obtenu le

THÉORÈME 5. *Pour que* $R(\rho)$ $(0 \leqslant \rho < \infty)$ *soit fonction de corrélation d'un champ aléatoire homogène et isotrope il est nécessaire et suffisant qu'elle admette la représentation* (13), *où* $g(\lambda)$ *est une fonction non décroissante monotone bornée.*

Lorsque $m = 2$, *la formule* (13) *s'écrit tout simplement*

$$R(\rho) = \int_0^\infty J_0(\lambda\rho)\, dg(\lambda), \tag{14}$$

et pour $m = 3$

$$R(\rho) = 2\int_0^\infty \frac{\sin \lambda\rho}{\lambda\rho}\, dg(\lambda). \tag{15}$$

CHAPITRE II

AXIOMATIQUE DE LA THÉORIE DES PROBABILITÉS

§ 1. Axiomes et définitions fondamentales de la théorie des probabilités

Dans le présent paragraphe on introduit des notions fondamentales de la théorie des probabilités et on donne, la plupart du temps sans les démontrer, leurs plus importantes propriétés, qui découlent directement de la théorie de la mesure. Les démonstrations se trouvent dans les manuels de théorie de la mesure (cf. par exemple A. Kolmogorov et S. Fomine [1]).

Evénements. Supposons que l'on réalise des expériences sous des conditions C. Ces expériences seront dites *possibles* (relativement aux conditions C). Les résultats d'une expérience nous permettent de conclure à la réalisation (ou non-réalisation) de tel ou tel événement. En théorie des probabilités chaque expérience est entièrement caractérisée par ces événements ou, plus exactement, par un ensemble non vide d'événements dits *observables*.

Pour rendre l'exposé plus clair on se servira de l'interprétation ensembliste des notions et relations introduites. Chaque événement observé à l'issue d'un possible est identifié à une partie d'un ensemble Ω. Les opérations et relations liant les événements admettent une interprétation ensembliste. Si par exemple l'événement A implique l'événement B ($A \subset B$), l'ensemble A sera contenu dans l'ensemble B. L'événement « ou A ou B » équivaut à la réunion des ensembles A et B ($A \cup B$), l'événement « A et B » à l'intersection des ensembles A et B ($A \cap B$), l'événement « non A » ($\bar{A}$) au complémentaire de A dans Ω, c'est-à-dire l'ensemble de tous les points de Ω n'appartenant pas à A, enfin, si deux événements A et B sont incompatibles, l'ensemble $A \cap B$ est vide.

Toute partie de Ω est à son tour un *événement*, mais qui n'est pas forcément réalisation d'une des expériences considérées.

Les points de Ω sont appelés *événements élémentaires* ou *éventualités*, l'ensemble Ω l'*événement certain* pour autant qu'il correspond à un événement réalisé dans toute expérience. Une partie vide de Ω est dite *événement impossible*.

On se servira encore des notations et définitions suivantes:

$\bigcup_{\alpha \in I} A_\alpha$: réunion de l'ensemble d'événements A_α indexés par $\alpha \in I$.

$\bigcap_{\alpha \in I} A_\alpha$: intersection de l'ensemble d'événements A_α.

$A \setminus B$: différence des événements A et B (l'événement « A mais non B »).

$A \Delta B$: différence symétrique des événements A et B (l'événement « ou A ou B, mais pas les deux à la fois »).

$\overline{\lim}\, A_n$: limite supérieure de la suite d'événements A_n ($n = 1, 2, \ldots$), événement qui est réalisé si et seulement si la suite d'événements A_n ($\overline{\lim}\, A_n = \{\text{ens. inf. } A_n\}$) l'est.

$\underline{\lim}\, A_n$: limite inférieure de la suite d'événements A_n ($n = 1, 2, \ldots$), événement réalisé si et seulement si tous les événements A_n à partir d'un certain numéro le sont. On remarquera sans peine que

$$\overline{\lim}\, A_n = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n, \quad \underline{\lim}\, A_n = \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n. \tag{1}$$

Par exemple, l'événement du second membre de la première égalité est réalisé si et seulement si, quel que soit m, on peut exhiber un $n \geqslant m$ tel que l'événement A_n soit réalisé. Il revient au même de dire qu'il existe une suite partielle infinie d'événements A_n réalisables.

Dans la suite on utilisera souvent les formules suivantes qui expriment une relation de dualité entre la réunion et l'intersection d'événements:

$$A \setminus \bigcup_{\alpha \in I} B_\alpha = \bigcap_{\alpha \in I} (A \setminus B_\alpha) \quad \Big(\text{ou } \overline{\bigcup_{\alpha \in I} B_\alpha} = \bigcap_{\alpha \in I} \overline{B}_\alpha\Big), \tag{2}$$

$$A \setminus \bigcap_{\alpha \in I} B_\alpha = \bigcup_{\alpha \in I} (A \setminus B_\alpha) \quad \Big(\text{ou } \overline{\bigcap_{\alpha \in I} B_\alpha} = \bigcup_{\alpha \in I} \overline{B}_\alpha\Big). \tag{3}$$

La vérification de ces formules ne soulève aucune difficulté d'autant plus que l'une d'elles est une conséquence de l'autre.

Expérience. Nous avons vu plus haut que, sous des conditions C données, tout possible était entièrement décrit par une classe (une famille) d'observables. On conviendra de noter les classes d'événements (classes de parties de l'ensemble Ω) par les lettres gothiques $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{F}$, ..., et par $\mathfrak{P}(\Omega)$ l'ensemble des parties de Ω.

Soit $\mathfrak{A}$ classe d'observables.

DÉFINITION. 1. *Une classe $\mathfrak{A} \subset \mathfrak{P}(\Omega)$ est un anneau booléen (ou clan) de parties de Ω si:*

$$A \in \mathfrak{A}, \quad B \in \mathfrak{A} \Rightarrow A \cup B \in \mathfrak{A},$$
$$A \in \mathfrak{A}, \quad B \in \mathfrak{A} \Rightarrow A \setminus B \in \mathfrak{A}.$$

Une classe $\mathfrak{A}$ est une algèbre booléenne (ou clan unitaire, ou anneau booléen unitaire) si $\mathfrak{A}$ est un anneau booléen et si $\Omega \in \mathfrak{A}$.

On démontre sans peine que:
si $\mathfrak{A}$ est anneau booléen: $A \in \mathfrak{A}$, $B \in \mathfrak{A} \Rightarrow A \cap B \in \mathfrak{A}$;
si $\mathfrak{A}$ est algèbre booléenne: $A \in \mathfrak{A} \Rightarrow \overline{A} \in \mathfrak{A}$.

DÉFINITION 2. *Une algèbre booléenne* $\mathfrak{A} \subset \mathfrak{B}(\Omega)$ *est une tribu* (*ou* *σ-algèbre, ou corps de Borel, ou σ-clan unitaire*) *si* $\forall\, n \in \mathrm{N}$

$$A_n \in \mathfrak{A} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathfrak{A}$$

On démontre sans peine que:

$$\forall n \in \mathrm{N} \quad A_n \in \mathfrak{A} \Rightarrow \bigcap_{n=1}^{\infty} A_n \in \mathfrak{A}.$$

En effet, ceci découle de la formule (3):

$$\bigcap_{n=1}^{\infty} A_n = \overline{\bigcup_{n=1}^{\infty} \overline{A}_n}.$$

Les événements de toute tribu $\mathfrak{A}$ sont dits $\mathfrak{A}$-mesurables.

On conviendra qu'une classe d'observables est une tribu.

L'ensemble des expériences peut être ordonné. Plus exactement, si des expériences E_1 et E_2 sont attachées à des tribus $\mathfrak{F}_1$ et $\mathfrak{F}_2$ d'observables et que $\mathfrak{F}_1 \subset \mathfrak{F}_2$, on dira que l'expérience E_2 est *plus informative* que E_1. D'autre part, si est donné un ensemble d'expériences $\{E_\alpha,\ \alpha \in I\}$ et $\mathfrak{F}_\alpha$ $(\alpha \in I)$ sont les tribus correspondantes d'observables, l'ensemble $\{E_\alpha,\ \alpha \in I\}$ peut être traité comme une expérience composée, à condition d'admettre que les observables appartiennent à la plus petite tribu contenant tous les $\mathfrak{F}_\alpha$. On notera cette dernière: $\sigma\,\{\mathfrak{F}_\alpha,\ \alpha \in I\}$. Son existence est une conséquence de la proposition suivante:

Soit $\mathfrak{M}$ *une classe de parties de* Ω. *Il existe toujours une plus petite tribu de parties contenant* $\mathfrak{M}$.

En effet, il existe toujours une tribu contenant $\mathfrak{M}$. Tel est par exemple l'ensemble $\mathfrak{B}(\Omega)$ de parties de Ω qui est une tribu. L'intersection de toutes les tribus contenant $\mathfrak{M}$ est une tribu: par construction, la plus petite contenant $\mathfrak{M}$. On la note $\sigma\,\{\mathfrak{M}\}$ et on l'appelle *tribu engendrée par la classe* $\mathfrak{M}$.

On aura souvent à établir si une classe d'événements contient ou non une tribu donnée. Le résultat suivant nous sera très utile.

THÉORÈME 1. *Si une classe d'ensembles* $\mathfrak{M}$ *contient une algèbre* $\mathfrak{A}$ *et est monotone, c'est-à-dire si* A_n, *suite monotone croissante* (*respectivement décroissante*) *d'ensembles et* $\bigcup_{n=1}^{\infty} A_n$ ($\bigcap_{n=1}^{\infty} A_n$) *sont contenus dans* $\mathfrak{M}$, *alors* $\mathfrak{M}$ *contient la tribu* $\sigma\{\mathfrak{A}\}$.

Probabilité. Soit $\mathfrak{S}$ la tribu de tous les observables dans un ensemble donné de possibles. La probabilité $\mathsf{P}(A)$ de l'événement $A \in \mathfrak{S}$ caractérise le lien existant entre les conditions C et les événements de $\mathfrak{S}$ quelle que soit l'expérience réalisée. L'interprétation naturelle de ce lien est suffisamment bien connue pour être rappelée. On introduira la notion axiomatique de la probabilité de la façon suivante (A. Kolmogorov [2], [7]).

Définition. *On appelle probabilité* $\mathsf{P}(\cdot)$ *une fonction numérique définie sur* $\mathfrak{S}$ *et possédant les propriétés suivantes:*

a) $$\mathsf{P}(A) \geqslant 0, \quad \mathsf{P}(\Omega) = 1,$$

b) $$\mathsf{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathsf{P}(A_n)$$

quelle que soit la suite d'événements deux à deux disjoints A_n, $n = 1, 2, \ldots$ ($A_n \cap A_r = \varnothing$ *pour* $n \neq r$, $A_n \in \mathfrak{S}$).

Donc la fonction P est une mesure définie sur $\mathfrak{S}$ et telle que $\mathsf{P}(\Omega) = 1$ (mesure probabiliste).

Définition. *Le triplet* $(\Omega, \mathfrak{S}, \mathsf{P}\}$ *constitué de l'ensemble* Ω, *de* $\mathfrak{S}$, *tribu d'observables, et de* P, *mesure probabiliste sur* $\mathfrak{S}$, *s'appelle espace probabilisé ou espace des probabilités.*

La probabilité possède les propriétés suivantes:

1. $\mathsf{P}(\bar{A}) = 1 - \mathsf{P}(A)$. En particulier, $\mathsf{P}(\varnothing) = 0$.

2. Si $A \subset B$, alors $\mathsf{P}(B \setminus A) = \mathsf{P}(B) - \mathsf{P}(A)$. En particulier, si $A \subset B$, alors $\mathsf{P}(A) \leqslant \mathsf{P}(B)$.

3. Si $A_n \subset A_{n+1}$, $n = 1, 2, \ldots$, alors

$$\mathsf{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim \mathsf{P}(A_n). \tag{4}$$

4. Si $A_n \supset A_{n+1}$, $n = 1, 2, \ldots$, alors

$$\mathsf{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim \mathsf{P}(A_n). \tag{5}$$

En particulier, si $A_n \supset A_{n+1}$, $\bigcap_{n=1}^{\infty} A_n = \varnothing$, alors

$$\mathsf{P}(A_n) \to 0. \tag{6}$$

Ces propriétés sont classiques en théorie de la mesure.

Dans bien des cas on aura besoin d'élargir le domaine de définition de la probabilité P à l'aide de l'opération suivante appelée complétion de l'espace probabilisé.

Soient $\mathfrak{N}$ classe des observables N de probabilité nulle, $\mathfrak{N} = \{N : \mathsf{P}(N) = 0,\ N \in \mathfrak{S}\}$; $\widetilde{\mathfrak{N}}$ classe des événements A tels qu'il existe un $N \in \mathfrak{N}$ tel que $A \subset N$, $\widetilde{\mathfrak{N}} = \{A : A \subset N,\ N \in \mathfrak{N}\}$; $\widetilde{\mathfrak{S}}$ classe des événements $\widetilde{S}$ ($\widetilde{S} = S \cup A$, $S \in \mathfrak{S}$, $A \in \widetilde{\mathfrak{N}}$). On vérifie sans peine que $\widetilde{\mathfrak{S}}$ est une tribu.

Définissons sur $\widetilde{\mathfrak{S}}$ la probabilité $\widetilde{\mathsf{P}}$ en posant $\widetilde{\mathsf{P}}(S \cup A) = \mathsf{P}(S)$ si $S \in \mathfrak{S}$ et $A \in \widetilde{\mathfrak{N}}$. Cette définition est univoque et $\widetilde{\mathsf{P}}$ est une fonction dénombrablement additive sur $\widetilde{\mathfrak{S}}$, donc $\{\Omega, \widetilde{\mathfrak{S}}, \widetilde{\mathsf{P}}\}$ est un espace probabilisé.

Le passage de $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ à $\{\Omega, \widetilde{\mathfrak{S}}, \widetilde{\mathsf{P}}\}$ s'appelle *complétion de l'espace probabilisé*. Si $\widetilde{\mathfrak{N}} \subset \mathfrak{N}$, alors $\widetilde{\mathfrak{S}} = \mathfrak{S}$ et l'espace probabilisé est par définition *complet*. De toute évidence $\{\Omega, \widetilde{\mathfrak{S}}, \widetilde{\mathsf{P}}\}$ est un espace probabilisé complet. Dans un tel espace toute partie d'un observable de probabilité nulle est elle-même observable.

Eléments aléatoires. Parfois la classe d'observables est donnée de la façon suivante. On postule que le résultat d'une expérience est décrit par un point d'un ensemble X. Si par exemple l'expérience consiste à mesurer la vitesse du vent à un instant donné et en un point donné de l'espace, pour X on peut prendre un espace vectoriel à trois dimensions. Si la vitesse du vent est constamment mesurée sur un intervalle de temps $[t_0, t_1]$ et si la fonction correspondante est supposée continue, pour X on peut prendre l'espace des fonctions continues sur $[t_0, t_1]$ et à valeurs dans un espace vectoriel à trois dimensions.

DÉFINITION. *On appelle espace probabilisable ou mesurable le couple* $\{X, \mathfrak{B}\}$, *où* X *est un ensemble et* $\mathfrak{B}$ *une tribu de parties de* X.

Désignons par ζ un point x caractérisant le résultat d'une expérience. Supposons que les observables sont de la forme $\{\zeta \in B\}$, $B \in \mathfrak{B}$. De ce qui précède il suit que l'événement $\{\zeta \in B\}$ est associé dans Ω à un ensemble $S = S_B \in \mathfrak{S}$. L'expérience considérée définit donc une application g de la tribu $\mathfrak{B}$ dans la tribu $\mathfrak{S}$. Cette application possède les propriétés suivantes:

a) $g(X) = \Omega$;

b) si $B_1 \cap B_2 = \varnothing$, $B_i \in \mathfrak{B}$, alors

$$g(B_1) \cap g(B_2) = \varnothing;$$

c) $g\left(\bigcup_{\alpha \in I} B_\alpha\right) = \bigcup_{\alpha \in I} g(B_\alpha)$, où I est un ensemble arbitraire d'indices, $B_\alpha \in \mathfrak{B}$.

De a), b) et c) il suit facilement que

d) $g(\varnothing) = \varnothing$;

e) $g(\overline{B}) = \overline{g(B)}$;

f) $g(B_2 \setminus B_1) = g(B_2) \setminus g(B_1)$;
g) $g(\bigcap_{\alpha\in I} B_\alpha) = \bigcap_{\alpha\in I} g(B_\alpha)$.

On dira qu'une application g de $\mathfrak{B}$ dans $\mathfrak{S}$ possédant les propriétés a, b) et c) engendre un élément aléatoire ζ dans $\{X, \mathfrak{B}\}$.

Posons $m(B) = \mathsf{P}(S)$ $(B \in \mathfrak{B})$. Il est évident que $m(B)$ est une mesure probabiliste et $\{X, \mathfrak{B}, m\}$ un espace probabilisé. La mesure m est appelée *répartition de l'élément aléatoire* ζ. De plus $m(B) = \mathsf{P}\{\zeta \in B\}$. Dans certains cas il est plus commode d'utiliser la notation

$$m = \mathsf{P}g \quad (\mathsf{P}g(B) = \mathsf{P}(g(B))).$$

Si $\mathfrak{B}$ contient des ensembles à un seul point $\{x\}$, $x \in X$, l'application g se laisse définir de la manière suivante.

Soit $g\{x\} = S_x$. En faisant parcourir l'ensemble X à x on partitionne l'espace Ω en une famille d'ensembles $\mathfrak{S}$-mesurables $\{S_x, x \in X\}$ deux à deux disjoints:

$$S_{x_1} \cap S_{x_2} = \varnothing \text{ pour } x_1 \neq x_2, \quad \bigcup_{x\in X} S_x = \Omega.$$

Soit une application f de Ω dans X, $f(\omega) = x$ si $\omega \in S_x$. De plus $S_B = g(B) = \{\omega : f(\omega) \in B\}$. Donc $g(B)$ est l'antécédent de l'ensemble B par l'application f de Ω dans X, $g(B) = f^{-1}(B)$. Soit par ailleurs f application de Ω dans X. L'application $f^{-1}(B)$ possède les propriétés a), b) et c), et si quel que soit $B \in \mathfrak{B}$

$$f^{-1}(B) = \{\omega : f(\omega) \in B\} \in \mathfrak{S}, \tag{7}$$

alors f^{-1} est un élément aléatoire de $\{X, \mathfrak{B}\}$.

DÉFINITION. *Une application f de* Ω *dans X vérifiant* (7) *s'appelle application mesurable de* $\{\Omega, \mathfrak{S}\}$ *dans* $\{X, \mathfrak{B}\}$.

De ce qui précède il suit que toute application ponctuelle mesurable f de $\{\Omega, \mathfrak{S}\}$ dans $\{X, \mathfrak{B}\}$ définit un élément aléatoire $\zeta = f(\omega)$ dans $\{X, \mathfrak{B}\}$. Inversement, si des ensembles X à un seul point sont $\mathfrak{B}$-mesurables, tout élément aléatoire de $\{X, \mathfrak{B}\}$ est défini grâce à l'application ponctuelle mesurable $f : \Omega \to X$. En particulier, ceci a lieu si X est un espace métrique, $\mathfrak{B}$ la tribu de ses boréliens.

La proposition suivante simplifie souvent la vérification de la réalisation de la condition de mesurabilité (7).

THÉORÈME 2. *Soit* $\mathfrak{B} = \sigma\{\mathfrak{M}\}$. *Pour que f soit application mesurable de* $\{\Omega, \mathfrak{S}\}$ *dans* $\{X, \mathfrak{B}\}$ *il suffit que la condition* (7) *soit réalisée pour tout* $B \in \mathfrak{M}$.

En effet, la classe des ensembles pour lesquels (7) a lieu est une tribu. Donc si (7) est réalisée pour tous les $B \in \mathfrak{M}$, elle l'est pour tous les $B \in \sigma\{\mathfrak{M}\}$. ■

Soit ζ élément aléatoire de $\{X, \mathfrak{B}\}$. Les propriétés a), b) et c) de l'application g entraînent que la classe d'événements

$$\sigma\{\zeta\} \underset{\text{Def}}{=} \{S : S = g(B),\ B \in \mathfrak{B}\} \tag{8}$$

est une tribu. On l'appelle *tribu engendrée par l'élément aléatoire* ζ. C'est une classe d'observables dans une expérience dont les issues possibles sont décrites par l'élément aléatoire ζ.

Dans la suite on envisagera essentiellement des éléments aléatoires définis par des applications ponctuelles de Ω dans X, c'est-à-dire des éléments $\zeta = f(\omega)$.

Soit donnée la suite d'événements aléatoires $\zeta_k = f_k(\omega)$, $k = 1, 2, \ldots, n$, à valeurs respectivement dans $\{X_k, \mathfrak{B}_k\}$. Cette suite peut être traitée comme un élément aléatoire ζ à valeurs dans un espace probabilisable $\{Y, \mathfrak{B}\}$ défini de la sorte. Soit Y l'ensemble des suites ordonnées $y = (x_1, x_2, \ldots, x_n)$, où $x_k \in X_k$. L'espace Y s'appelle *produit des espaces* $X_1, \ldots, X_n$ et se note $Y = \prod_{k=1}^{n} X_k$ ou encore $Y = X_1 \times X_2 \times \ldots \times X_n$. Considérons dans Y la classe d'ensembles B de la forme $B = \prod_{k=1}^{n} B_k$, $B_k \in \mathfrak{B}_k$, c'est-à-dire

$$B = \{y = (x_1, x_2, \ldots, x_n) : x_k \in B_k,\ k = 1, \ldots, n\}.$$

De tels ensembles sont appelés *pavés* dans Y. La plus petite tribu $\mathfrak{B}$ contenant tous les pavés est appelée *produit des tribus* $\mathfrak{B}_k$ et s'écrit $\mathfrak{B} = \sigma\{\mathfrak{B}_k, k = 1, 2, \ldots, n\}$ et l'espace probabilisable $\{Y, \mathfrak{B}\} = \coprod_{k=1}^{n} \{X_k, \mathfrak{B}_k\}$ *produit cartésien des espaces* $\{X_k, \mathfrak{B}_k\}$.

Soit f application de Ω dans Y, définie par $Y = f(\omega) = (f_1(\omega), f_2(\omega), \ldots, f_n(\omega))$. Si $B = \coprod_{k=1}^{n} B_k$, alors $f^{-1}(B) = \bigcap_{k=1}^{n} f^{-1}(B_k) \in \mathfrak{S}$.

Toute classe $\mathfrak{A}$ d'ensembles A tels que $f^{-1}(A) \in \mathfrak{S}$ est une tribu (vu que l'antécédent de la réunion, de l'intersection et de la différence d'ensembles est égal respectivement à la réunion, l'intersection et la différence des antécédents). Comme $\mathfrak{A}$ contient les pavés, elle contient la plus petite tribu $\mathfrak{B}$ engendrée par ces pavés. Donc f est une application mesurable de $\{\Omega, \mathfrak{S}\}$ dans $\{Y, \mathfrak{B}\}$. On dira que l'élément aléatoire $\zeta = f(\omega)$ est le produit cartésien des éléments aléatoires $\zeta_k = f_k(\omega)$ $(k = 1, 2, \ldots, n)$.

On appelle *variable aléatoire* un élément aléatoire $\zeta = f(\omega)$ à valeurs réelles ($X = \mathcal{R}^1$, $\mathfrak{B} = \mathfrak{B}^1$ est la tribu des boréliens de l'axe des réels). On appelle *vecteur aléatoire* un élément aléatoire ζ

à valeurs dans l'espace réel $\mathscr{R}^n$ de n dimensions ($\mathfrak{B} = \mathfrak{B}^n$ est la tribu des boréliens de l'espace à n dimensions).

Variables aléatoires. Une variable aléatoire ζ est définie par une fonction réelle $f(\omega)$ telle que $f^{-1}(B) \in \mathfrak{S}$ pour tout borélien $B \in \mathscr{R}^1$. Comme la tribu des boréliens de la droite est engendrée par un système d'intervalles infinis $\{]-\infty, a[, a \in \mathscr{R}^1\}$, pour que f soit mesurable il suffit que, quel que soit a, l'on ait $f^{-1}(-\infty, a) = \{\omega: f(\omega) < a\} \in \mathfrak{S}$. Cette dernière condition figure généralement dans la définition de toute fonction réelle $\mathfrak{S}$-mesurable définie sur l'espace probabilisable $\{\Omega, \mathfrak{S}\}$. Donc la notion de variable aléatoire coïncide avec celle de fonction réelle $\mathfrak{S}$-mesurable et les propriétés générales des variables aléatoires sont celles des fonctions mesurables réelles.

Remarque. Parfois on a à considérer des éléments aléatoires à valeurs dans la droite numérique achevée $\widetilde{\mathscr{R}}^1 = [-\infty, +\infty]$. De tels éléments aléatoires sont appelés *variables aléatoires généralisées.* La tribu $\mathfrak{B}$ est ici la tribu $\widetilde{\mathfrak{B}}^1$ composée d'ensembles de la forme

$$B,\ B \cup \{-\infty\},\ B \cup \{+\infty\},\ B \cup \{+\infty\} \cup \{-\infty\} \quad (B \in \mathfrak{B}^1).$$

Voyons quelques propriétés des variables aléatoires.

Théorème 3. *Une fonction borélienne* $\eta = g(\xi_1, \ldots, \xi_n) = g(f_1(\omega), \ldots, f_n(\omega))$ *de variables aléatoires* $\xi_k = f_k(\omega)$, $k = 1, \ldots, n$, *est aussi variable aléatoire et* ξ_1/ξ_2, $\sup \xi_n$, $\inf \xi_n$, $\overline{\lim}\ \xi_n$, $\underline{\lim}\ \xi_n$ *sont des variables aléatoires généralisées.*

Le rapport de deux variables aléatoires est par convention nul si le numérateur et le dénominateur le sont simultanément.

Les propriétés indiquées des variables aléatoires (fonctions mesurables) sont classiques en la théorie de la mesure.

Les indicateurs d'événements mesurables constituent un important exemple de variables aléatoires. L'*indicateur* $\chi(A) = \chi(A, \omega)$ se définit comme suit:

Définition.

$$\chi(A, \omega) = \begin{cases} 1 & si\ \ \omega \in A, \\ 0 & si\ \ \omega \notin A. \end{cases}$$

On peut effectuer les mêmes opérations algébriques sur les indicateurs que sur les événements. En effet,

$$\chi\left(\bigcap_1^\infty A_n\right) = \prod_1^\infty \chi(A_n),$$

$$\chi\left(\bigcup_1^\infty A_n\right) = \sum_1^\infty \chi(A_n), \text{ si } A_n \cap A_r = \varnothing, \quad n \neq r,$$

$$\chi(A \setminus B) = \chi(A) - \chi(B), \text{ si } B \subset A,$$

$$\chi(\overline{\lim} A_n) = \overline{\lim}\,\chi(A_n),$$
$$\chi(\underline{\lim} A_n) = \underline{\lim}\,\chi(A_n).$$

Une variable aléatoire prenant seulement un nombre fini ou dénombrable de valeurs est dite *discrète*. Si $\{c_n, n = 1, 2, \ldots\}$ est l'ensemble de valeurs possibles d'une variable aléatoire ξ, A_k l'événement $\{\xi = c_n\}$, alors $\xi = \sum\limits_n c_n \chi(A_n)$.

THÉORÈME 4. *Pour toute variable aléatoire ξ il existe une suite de variables aléatoires discrètes ξ_n ne prenant qu'un nombre fini de valeurs distinctes et convergeant vers ξ pour chaque ω. Si ξ est non négative, il existe une suite monotone non décroissante de variables aléatoires discrètes ξ'_n telles que $0 \leqslant \xi - \xi'_n \leqslant \frac{1}{2^n}$ quel que soit ω.*

En effet, si l'on pose

$$\xi_n = \sum_{j=-n}^{n-1} \sum_{k=1}^{n} \left(j + \frac{k-1}{n}\right) \chi(A_{jk}),$$

où

$$A_{jk} = \left\{\omega : j + \frac{k-1}{n} \leqslant \xi < j + \frac{k}{n}\right\},$$

alors pour $|\xi| < n$, on aura

$$0 \leqslant \xi - \xi_n \leqslant \frac{1}{n}.$$

En posant

$$\xi'_n = \sum_{k=0}^{\infty} \frac{k}{2^n} \chi(A'_{nk}),$$

où $A'_{nk} = \left\{\omega : \frac{k}{2^n} \leqslant \xi < \frac{k+1}{2^n}\right\}$, on obtient

$$0 \leqslant \xi - \xi'_n < \frac{1}{2^n}$$

pour tous les ω tels que $\xi \geqslant 0$. A noter que la suite ξ'_n est monotone non décroissante. ∎

Supposons que l'on puisse déterminer une variable aléatoire ξ d'après une expérience dont les issues sont décrites par un élément aléatoire ζ à valeurs dans un espace probabilisable $\{X, \mathfrak{B}\}$. Naturellement ξ sera une fonction $\mathfrak{B}$-mesurable de l'élément aléatoire ζ. La formulation exacte de cette proposition fait l'objet du

THÉORÈME 5. *Si ξ est une variable aléatoire mesurable par rapport à la tribu $\mathfrak{F}_\zeta$ engendrée par un élément aléatoire $\zeta = f(\omega)$ sur un*

espace probabilisable $\{X, \mathfrak{B}\}$, *il existe une fonction* $\mathfrak{B}$*-mesurable* $g(x)$, $x \in X$, *telle que* $\xi = g(\zeta)$.

Démonstration. Supposons tout d'abord que ξ prend un nombre fini ou dénombrable de valeurs a_n, $n = 1, 2, \ldots$ Posons $A_n = \{\omega : \xi = a_n\}$. ξ étant $\mathfrak{F}_\zeta$-mesurable, $A_n \in \mathfrak{F}_\zeta$ et par conséquent il existe un $B_n \in \mathfrak{B}$ tel que $f^{-1}(B_n) = A_n$. Soient $C_n = B_n \setminus \bigcup_{k=1}^{n-1} B_k$. Il vient

$$C_n \in \mathfrak{B}, \quad C_n \cap C_m = \varnothing \text{ pour } n \neq m,$$

$$f^{-1}(C_n) = f^{-1}(B_n) \setminus \bigcap_1^{n-1} f^{-1}(B_k) = A_n \setminus \bigcup_1^{n-1} A_k = A_n$$

et

$$f^{-1}\left(\bigcup_1^\infty C_n\right) = \bigcup_1^\infty A_n = \Omega,$$

c'est-à-dire

$$f(\Omega) \subset \bigcup_1^\infty C_n.$$

Posons $g(x) = \sum a_n \chi(C_n, x)$. Alors $\xi = g(\zeta) = g(f(\omega))$. Généralisons. Il existe une suite de variables aléatoires discrètes ξ_n, $\mathfrak{F}_\zeta$-mesurables et convergentes vers ξ pour tout ω. D'après ce qui précède $\xi_n = g_n(\zeta)$, où $g_n(x)$ est une fonction $\mathfrak{B}$-mesurable. L'ensemble S des points de convergence de $g_n(x)$ est $\mathfrak{B}$-mesurable, contient $f(\Omega)$ et pour tout $x = f(\omega) = \zeta$ existe $\lim g_n(x) = \xi$. Soit

$$g(x) = \begin{cases} \lim g_n(x) & \text{pour } x \in S, \\ 0 & \text{pour } x \notin S. \end{cases}$$

Alors $g(x)$ est une fonction $\mathfrak{B}$-mesurable et $g(\zeta) = \xi$. ■

Définition. *Une tribu* $\mathfrak{F}$ *est simple si elle est composée de* $\varnothing$ *et de tous les ensembles* $F = \sum_j E_{n_j}$, *où* $\{E_n, n = 1, 2, \ldots\}$ *est une suite dénombrable (ou finie) d'événements deux à deux disjoints et* $\bigcup_n E_n = \Omega$. *Les ensembles* E_n *seront appelés atomes de la tribu* $\mathfrak{F}$.

Théorème 6. *Si une variable aléatoire* η *est mesurable par rapport à une tribu simple* $\mathfrak{F}$, *elle est constante sur les atomes de* $\mathfrak{F}$:

$$\eta = \sum_n a_n \chi(E_n).$$

En effet, la variable aléatoire η est susceptible d'être approchée par une suite, convergente pour tout ω, de variables aléatoires η_n, constantes sur les atomes E_n (théorème 4). D'où il suit que η est constante sur les ensembles E_n. ■

On dit qu'une proposition ou une propriété attachée à un espace probabilisé $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ a lieu avec la *probabilité* 1 ou est *presque sûre* si l'événement qui consiste en ce qu'elle se réalise est $\mathfrak{S}$-mesurable et possède une probabilité égale à 1. Si $\xi = \eta$ presque sûrement, les variables aléatoires ξ et η sont dites *équivalentes*, ce qu'on note $\xi = \eta \pmod{\mathsf{P}}$.

THÉORÈME 7. *Si* $\xi_n = \eta_n \pmod{\mathsf{P}}$, $n = 1, 2, \ldots$, *et* $h(t_1, \ldots, t_n)$ *est une fonction borélienne sur* $\mathcal{R}^n$, *les relations suivantes ont presque sûrement lieu :*

$$h(\xi_1, \ldots, \xi_n) = h(\eta_1, \ldots, \eta_n), \quad \sup \xi_n = \sup \eta_n,$$
$$\inf \xi_n = \inf \eta_n, \quad \overline{\lim}\, \xi_n = \overline{\lim}\, \eta_n, \quad \underline{\lim}\, \xi_n = \underline{\lim}\, \eta_n.$$

Cette proposition montre que généralement les opérations analytiques transforment les systèmes de variables aléatoires équivalentes en systèmes équivalents de sorte qu'on peut parler d'opérations sur des classes de variables équivalentes.

Convergence presque sûre. Soit ξ_n, $n = 1, 2, \ldots$, une suite de variables aléatoires. L'événement $S = \{\lim \xi_n \text{ existe}\}$ est $\mathfrak{S}$-mesurable.

En effet

$$S = \bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m_1, m_2 \geqslant n} \left\{\omega : |\xi_{m_1} - \xi_{m_2}| < \frac{1}{k}\right\}.$$

DÉFINITION. *Une suite* ξ_n, $n = 1, 2, \ldots$, *converge presque sûrement ou avec la probabilité* 1 *si*

$$\mathsf{P}\{\lim \xi_n \text{ existe}\} = 1.$$

Pour démontrer la convergence presque sûre dans des problèmes concrets on aura besoin du critère suivant.

THÉORÈME 8. *Si existe une suite de nombres* ε_n *tels que*

$$\varepsilon_n > 0, \quad \sum_{n=1}^{\infty} \varepsilon_n < \infty, \quad \sum_{n=1}^{\infty} \mathsf{P}\{|\xi_{n+1} - \xi_n| > \varepsilon_n\} < \infty,$$

alors $\lim \xi_n$ *existe presque sûrement.*

Si pour tout $\varepsilon > 0$

$$\sum_{n=1}^{\infty} \mathsf{P}\{|\xi - \xi_n| > \varepsilon\} < \infty,$$

alors ξ_n *converge presque sûrement vers* ξ.

D é m o n s t r a t i o n . Soit

$$A_n = \{\omega : |\xi_{n+1} - \xi_n| > \varepsilon_n\}.$$

Il vient

$$\mathsf{P}\,(\overline{\lim}\, A_n) = \mathsf{P}\,(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m) = \lim \mathsf{P}\,(\bigcup_{n} A_m) \leqslant \overline{\lim} \sum_{n}^{\infty} \varepsilon_m = 0.$$

Donc il existe presque sûrement un $n_0 = n_0(\omega)$ tel que $|\xi_{n+1} - \xi_n| \leqslant \varepsilon_n$ pour tous les $n \geqslant n_0$. D'où il suit que la série $\xi_1 + \sum_{n=1}^{\infty} (\xi_{n+1} - \xi_n)$ est presque sûrement convergente, ce qui démontre la première assertion. Supposons maintenant que $A_{N_n} = \left\{\omega : |\xi - \xi_n| > \frac{1}{N}\right\}$.

Il vient

$$\mathsf{P}\,\{\underline{\lim}\, |\xi - \xi_n| > 0\} = \mathsf{P}\,\{\bigcup_{N=1}^{\infty} \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_{N_n}\} \leqslant$$

$$\leqslant \lim_{N\to\infty} \lim_{m\to\infty} \sum_{n=m}^{\infty} \mathsf{P}\,(A_{N_m}) = 0. \quad \blacksquare$$

Convergence en probabilité (ou stochastique).

DÉFINITION. *Une suite de variables aléatoires* ξ_n, $n = 1, 2, \ldots$, *converge stochastiquement (ou en probabilité) vers une variable aléatoire* ξ (P-lim $\xi_n = \xi$) *si pour tout* $\varepsilon > 0$

$$\mathsf{P}\,\{|\xi_n - \xi| > \varepsilon\} \to 0 \textit{ lorsque } n \to \infty.$$

La suite ξ_n *est fondamentale stochastiquement si pour tout* $\varepsilon > 0$ *on peut exhiber un* $n_0 = n_0(\varepsilon)$ *tel que pour tous les* $n_1, n_2 \geqslant n_0$

$$\mathsf{P}\{|\xi_{n_1} - \xi_{n_2}| > \varepsilon\} < \varepsilon.$$

La notion de convergence stochastique correspond à la notion de convergence en mesure dans la théorie générale de la mesure. De cette théorie on déduit le

THÉORÈME 9. a) *Si* $\eta_i = \mathsf{P}$-lim ξ_n, $i = 1, 2$, *alors* $\eta_1 = \eta_2$ (mod P).

b) *Pour qu'une suite de variables aléatoires* ξ_n *converge stochastiquement il faut et il suffit qu'elle soit fondamentale stochastiquement.*

c) *Si une suite* ξ_n *est convergente presque sûrement, elle l'est stochastiquement. En général, la réciproque n'est pas vraie. Mais d'une suite de variables aléatoires convergente stochastiquement on peut extraire une suite partielle convergente presque sûrement.*

d) *Soient* $\eta^{(k)} = \mathsf{P}$-lim $\xi_n^{(k)}$ $(k = 1, \ldots, d)$, $\zeta_n = g(\xi_n^{(1)}, \ldots, \xi_n^{(d)})$, *où* $g(t_1, \ldots, t_d)$ *est une fonction réelle continue sur* $\mathscr{R}^d$ *sauf peut-être en un ensemble de points D tel que* $\mathsf{P}\{(\eta^{(1)}, \ldots, \eta^{(d)}) \in D\} = 0$. *On a alors*

$$\mathsf{P}\text{-lim}\, \zeta_n = g(\eta^{(1)}, \ldots, \eta^{(d)}).$$

En particulier, les suites $\{\xi_n^{(k)}\}$ convergent stochastiquement en même temps que les suites $\xi_n^{(1)} + \xi_n^{(2)}$, $\xi_n^{(1)} \cdot \xi_n^{(2)}$ et $\xi_n^{(1)}/\xi_n^{(2)}$, la dernière sous réserve que $\mathsf{P}\{\eta^{(2)} = 0\} = 0$, et

$$\mathsf{P}\text{-}\lim (\xi_n^{(1)} + \xi_n^{(2)}) = \eta^{(1)} + \eta^{(2)},$$
$$\mathsf{P}\text{-}\lim (\xi_n^{(1)} \cdot \xi_n^{(2)}) = \eta^{(1)} \cdot \eta^{(2)},$$
$$\mathsf{P}\text{-}\lim \frac{\xi_n^{(1)}}{\xi_n^{(2)}} = \frac{\eta^{(1)}}{\eta^{(2)}}.$$

Espérance mathématique. L'espérance mathématique d'une variable aléatoire discrète $\xi = \sum_n c_n \chi(A_n)$ est définie par

$$\mathsf{E}\xi = \sum_n c_n \mathsf{P}\{\xi = c_n\}$$

si la somme de la série a un sens. Cette expression est l'intégrale de la fonction simple $f(\omega) = \sum c_n \chi(A_n)$:

$$\mathsf{E}\xi = \int_\Omega f(\omega)\, \mathsf{P}(d\omega) \tag{9}$$

dans un espace mesuré abstrait $\{\Omega, \mathfrak{S}, \mathsf{P}\}$. On admettra que la dernière formule définit l'espérance mathématique dans le cas général sous réserve de l'existence de l'intégrale du second membre. Ceci signifie qu'au moins une des intégrales

$$\int_\Omega f^+(\omega)\, \mathsf{P}(d\omega), \quad \int_\Omega f^-(\omega)\, \mathsf{P}(d\omega)$$

est finie ($f^+ = \max\{f, 0\}$, $f^{-1} = \max\{-f, 0\}$). Si l'une d'elles est infinie, alors $\mathsf{E}\xi = +\infty$ ou $-\infty$ respectivement. Si l'espérance mathématique d'une variable ξ est finie, on dira que cette variable est *intégrable.*

On se servira également de la notation

$$\mathsf{E}\xi = \int_\Omega \xi\, d\mathsf{P}.$$

Posons

$$\int_A \xi\, d\mathsf{P} = \int_A f(\omega)\, \mathsf{P}(d\omega) = \mathsf{E}\chi(A)\,\xi.$$

Les propriétés générales de l'intégrale entraînent les propriétés suivantes de l'espérance mathématique.

THÉORÈME 10. a) *L'inégalité*

$$\int_A \xi\, d\mathsf{P} \geqslant \int_A \eta\, d\mathsf{P}$$

est réalisée pour tous les $A \in \mathfrak{S}$ *si et seulement si*

$$\xi \geqslant \eta \pmod{\mathsf{P}};$$

b)
$$\int_A \xi \, d\mathsf{P} = \int_A \eta \, d\mathsf{P}$$

pour tous les A *si et seulement si* $\xi = \eta \pmod{\mathsf{P}}$;

c) *si* $\mathsf{E}\xi$ *et* $\mathsf{E}\eta$ *sont finies, alors*

$$\mathsf{E}(a\xi + b\eta) = a\mathsf{E}\xi + b\mathsf{E}\eta$$

pour tous les a *et* b *constants.*

Signalons les inégalités suivantes d'un usage courant.

Inégalité de Tchébychev. *Si* $f(x) > 0$ $(x > 0)$ *est une fonction monotone non décroissante* $(x \geqslant 0)$, *alors* $\forall a > 0$

$$\mathsf{P}\{|\xi| > a\} \leqslant \frac{\mathsf{E}f(|\xi|)}{f(a)}.$$

Inégalité de Jensen. *Si* $g(x)$ *est une fonction convexe continue sur* $]a, b[$, $-\infty \leqslant a < b \leqslant +\infty$ *et* $a < \xi < b \pmod{\mathsf{P}}$, *alors*

$$g(\mathsf{E}\xi) \leqslant \mathsf{E}g(\xi). \tag{10}$$

On rappelle qu'une fonction $g(x)$ est *convexe sur* $]a, b[$ si

$$g(\lambda x_1 + (1 - \lambda) x_2) \leqslant \lambda g(x_1) + (1 - \lambda) g(x_2),$$

x_1, x_2 étant deux points quelconques de $]a, b[$ et quel que soit $\lambda \in$ $\in]0, 1[$, autrement dit si aucun point de la portion de courbe $y =$ $= g(x)$, $x \in]x_1, x_2[$ n'est situé au-dessus de la corde qui joint les extrémités de cette portion. La définition entraîne que pour tout point $Q = (x_0, g(x_0))$ du graphe d'une courbe convexe il existe une droite d'appui, c'est-à-dire une droite passant par Q et telle que tous les points de la courbe soient situés au-dessus d'elle. Ce qui revient à dire que pour tout $x_0 \in]a, b[$ il existe un α tel que

$$g(x) - g(x_0) \geqslant \alpha(x - x_0)$$

pour tous les $x \in]a, b[$. En posant $x_0 = \mathsf{E}\xi$, $x = \xi$ et en prenant l'espérance mathématique des deux membres de cette inégalité on obtient l'inégalité (10).

COROLLAIRE. $|\mathsf{E}\xi|^p \leqslant \mathsf{E}|\xi|^p$ *pour* $p > 1$ *et*

$$(\mathsf{E}|\xi|^q)^{\frac{1}{q}} \leqslant (\mathsf{E}|\xi|^p)^{\frac{1}{p}} \quad pour \quad 0 < q < p. \tag{11}$$

Inégalité de Hölder.

$$\mathsf{E}|\xi\eta| \leqslant (\mathsf{E}|\xi|^p)^{\frac{1}{p}} \cdot (\mathsf{E}|\eta|^q)^{\frac{1}{q}}, \quad où \ \frac{1}{p} + \frac{1}{q} = 1, \quad p > 1. \tag{12}$$

L'inégalité de Cauchy-Bouniakovski

$$(\mathsf{E}\xi\eta)^2 \leqslant \mathsf{E}\xi^2\cdot\mathsf{E}\eta^2$$

en est un cas particulier. L'inégalité de Hölder permet de déduire sans peine l'inégalité de Minkowski

$$[\mathsf{E}\,|\xi+\eta|^p]^{\frac{1}{p}} \leqslant [\mathsf{E}\,|\xi|^p]^{\frac{1}{p}} + [\mathsf{E}\,|\eta|^p]^{\frac{1}{p}}, \quad p\geqslant 1. \tag{13}$$

On utilise souvent le passage à la limite sous le signe de l'espérance mathématique.

THÉORÈME 11. a) Théorème de la convergence monotone. *Si* $0 \leqslant \xi_n \leqslant \xi_{n+1}$ $(n = 1, 2, \ldots)$, *alors*

$$\lim \mathsf{E}\xi_n = \mathsf{E}\,(\lim \xi_n).$$

b) Lemme de Fatou. *Si* $\xi_n \geqslant 0$ (mod P), *alors*

$$\mathsf{E}\,\underline{\lim}\,\xi_n \leqslant \underline{\lim}\,\mathsf{E}\,\xi_n.$$

c) Théorème de Lebesgue de la convergence majorée. *Si* $|\xi_n| \leqslant \eta$ (mod P), $\lim \xi_n = \xi$ (mod P) *et* $\mathsf{E}\eta < \infty$, *alors*

$$\lim \mathsf{E}\xi_n = \mathsf{E}\xi.$$

Posons

$$\varphi(A) = \int_A \xi\, d\mathsf{P} = \mathsf{E}\xi\chi(A), \quad A \in \mathfrak{S}. \tag{14}$$

DÉFINITION. *Une fonction réelle d'ensemble* $\varphi(A)$, *définie sur une tribu* $\mathfrak{S}$, *est une charge si elle peut prendre des valeurs infinies de même signe et si pour toute suite d'ensembles disjoints* A_n $(A_n \in \mathfrak{S})$, $A_n \cap A_r = \varnothing$ *pour* $n \neq r$, $n = 1, 2, \ldots)$

$$\varphi\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \varphi(A_n).$$

Si une variable aléatoire ξ possède une espérance mathématique (finie ou infinie), la fonction d'ensembles φ définie par (14) est une charge.

La formule suivante définit le principe de changement des variables dans la théorie de l'intégration.

THÉORÈME 12. *Soit* $\xi = g(\zeta)$, *où* $\zeta = f(\omega)$ *est un élément aléatoire de* $\{X, \mathfrak{B}\}$, *g une fonction réelle* $\mathfrak{B}$*-mesurable sur* X, $m = \mathsf{P}f^{-1}$ *la répartition de l'élément aléatoire* ζ. *Alors*

$$\mathsf{E}\xi = \int_X g(x)\, m(dx) \tag{15}$$

si l'un des deux membres a un sens.

Supposons qu'une mesure q finie ou σ -finie est donnée sur $\{X, \mathfrak{B}\}$. On rappelle qu'une mesure q est *σ-finie* si existe une suite d'ensembles $B_n \in \mathfrak{B}$ tels que $\bigcup_n B_n = X$ et $q(B_n) < \infty$.

DÉFINITION. *Une mesure m sur $\{X, \mathfrak{B}\}$ est absolument continue par rap ort à une mesure q ($m \ll q$) si existe une fonction $\mathfrak{B}$-mesurable $\rho(x)$ telle que*

$$m(B) = \int_B \rho(x)\, q(dx) \text{ pour tous les } B \in \mathfrak{B}. \tag{16}$$

THÉORÈME 13 (théorème de Radon-Nikodym). *Pour qu'une mesure m soit absolument continue par rapport à une mesure q il faut et il suffit que $q(B) = 0$ implique $m(B) = 0$.*

Soit $m = \mathsf{P}f^{-1}$ la répartition d'un élément aléatoire $\zeta = f(\omega)$ sur $\{X, \mathfrak{B}\}$ et $m \ll q$.

DÉFINITION. *Une fonction ρ définie par la relation* (16) *s'appelle densité de probabilité de l'élément aléatoire ζ par rapport à la mesure q. La densité de probabilité possède les propriétés suivantes*:

$$\rho(x) \geqslant 0 \ (\mathrm{mod}\ \mathsf{P}), \quad \int_X \rho(x)\, q(dx) = 1.$$

Si une répartition possède une densité, l'espérance mathématique de la variable $\xi = f(\zeta)$ se calcule à l'aide de la formule

$$\mathsf{E}\xi = \int_X f(x)\, \rho(x)\, q(dx). \tag{17}$$

L'espace $\mathscr{L}_p$. Soit $\mathscr{L}_p = \mathscr{L}_p(\Omega, \mathfrak{S}, \mathsf{P})$ la classe des variables aléatoires ξ telles que $\mathsf{E}\,|\xi|^p < \infty$ ($p > 0$). Cette classe est linéaire: si $\xi_i \in \mathscr{L}_p$ ($i = 1, 2$), alors $c\xi_i \in \mathscr{L}_p$ et $\xi_1 + \xi_2 \in \mathscr{L}_p$. La première assertion est évidente. La deuxième découle de l'inégalité de Minkowski (13) pour $p \geqslant 1$ et de l'inégalité

$$|\xi_1 + \xi_2|^p \leqslant |\xi_1|^p + |\xi_2|^p \tag{18}$$

pour $p \in]0, 1[$, laquelle dérive de $(1 + x)^p < 1 + x^p$ ($x > 0$, $0 < p < 1$).

Si l'on minit $\mathscr{L}_p$ ($p \geqslant 1$) de la norme

$$\|\xi\|_p = (\mathsf{E}\,|\xi|^p)^{\frac{1}{p}},$$

ce sera un espace vectoriel normé.

La convergence dans $\mathscr{L}_p$ d'une suite ξ_n vers ξ signifie que

$$\mathsf{E}\,|\xi - \xi_n|^p \to 0 \quad \text{lorsque} \quad n \to \infty.$$

La complétude de l'espace $\mathscr{L}_p$ entraîne qu'une condition nécessaire et suffisante pour qu'une suite ξ_n soit convergente est que soit réalisée la condition de Cauchy

$$\mathsf{E}\,|\xi_{n_1} - \xi_{n_2}|^p \to 0 \quad \text{lorsque} \quad n_1, n_2 \to \infty.$$

Toute suite de variables aléatoires satisfaisant cette condition est dite *fondamentale dans* $\mathcal{L}_p$. Une suite fondamentale (et par conséquent convergente) dans $\mathcal{L}_p$ est fondamentale (resp. convergente) stochastiquement. Ceci découle de l'inégalité de Tchébychev

$$\mathsf{P}\{|\xi_{n_1}-\xi_{n_2}|>\varepsilon\}\leqslant\frac{1}{\varepsilon^p}\mathsf{E}\,|\xi_{n_1}-\xi_{n_2}|^p.$$

Equi-intégrabilité. Une famille de variables aléatoires $\{\xi(t),\ t\in T\}$ (T est un ensemble) est *équi-intégrable* (ou *uniformément intégrable*) si pour tout $\varepsilon>0$ on peut exhiber un $N=N(\varepsilon)$ ne dépendant pas de t tel que

$$\mathsf{E}\chi(\{|\xi(t)|>N\})\,|\xi(t)|<\varepsilon\ \forall t\in T.$$

La proposition suivante donne des propriétés qui sont équivalentes à l'équi-intégrabilité.

THÉORÈME 14. *Une famille* $\{\xi(t),\ t\in T\}$ *est équi-intégrable si et seulement si*

a) $\mathsf{E}\,|\xi(t)|\leqslant C\ \forall t\in T$;

b) *pour tout* $\varepsilon>0$ *il existe un* $\delta>0$ *ne dépendant pas de t et tel que quel que soit A mesurable tel que* $\mathsf{P}(A)<\delta$, *on a* $\mathsf{E}\chi(A)\,|\xi(t)|<\varepsilon$.

On remarquera qu'une ou un nombre fini de variables aléatoires équi-intégrables forment une famille équi-intégrable et que l'union de deux ou d'un nombre fini de familles équi-intégrables est de nouveau une famille équi-intégrable.

Les critères suivants d'équi-intégrabilité nous seront utiles dans des cas concrets.

THÉORÈME 15. a) *Si une famille* $\{\xi(t),\ t\in T\}$ *est majorée par une variable aléatoire intégrable* η ($|\xi(t)|\leqslant\eta$ (mod P), $\mathsf{E}\eta<\infty$), *elle est équi-intégrable.*

b) *Soit* $g(x)$, $x\in\,]-\infty,\infty[$, *une fonction borélienne non négative telle que* $g(x)/x\to\infty$ *avec* $|x|$ *et*

$$\mathsf{E}g(\xi(t))\leqslant c\ \forall\ t\in T,$$

où c ne dépend pas de t. La famille $\{\xi(t),\ t\in T\}$ *est équi-intégrable.*

Les critères de convergence dans $\mathcal{L}_1$ s'énoncent simplement en termes d'équi-intégrabilité.

THÉORÈME 16. *Pour qu'une suite* $\{\xi_n,\ n=1,2,\ldots\}$ *soit convergente dans* $\mathcal{L}_1$ *il faut et il suffit qu'elle soit convergente stochastiquement et équi-intégrable.*

Les critères de convergence dans $\mathcal{L}_p$ ($p\geqslant 1$) se formulent de façon analogue. Au lieu de l'équi-intégrabilité de la suite ξ_n il faudra exiger alors celle de $\{|\xi_n|^p,\ n=1,2,\ldots\}$. Les démonstra-

tions de ces propositions figurent dans A. Kolmogorov et S. Fomine [1], J. Neveu [1], P. Hennequin et A. Tortrat [1].

§ 2. Construction d'espaces probabilisés

L'espace probabilisé est un objet mathématique assez complexe et dans nombre de problèmes on ne peut le considérer comme donné *a priori*. Aussi est-il important de savoir le construire. Dans les cas les plus simples il faut construire un espace probabilisé finidimensionnel d'après la fonction de répartition qui est donnée. Voyons quelques propriétés des fonctions de répartition.

Fonctions de répartition. Soit ξ vecteur aléatoire à valeurs dans $\mathcal{R}^d$, $\xi = (\xi^1, \xi^2, \ldots, \xi^d)$.

Posons

$$F(x) = F(x^1, x^2, \ldots, x^d) = \mathsf{P}\{\xi^1 < x^1, \xi^2 < x^2, \ldots, \xi^d < x^d\}.$$

DÉFINITION. *La fonction $F(x)$ s'appelle fonction de répartition du vecteur aléatoire ξ (ou fonction de répartition conjointe des variables aléatoires $\xi^1, \xi^2, \ldots, \xi^d$).*

Soit $a, b \in \mathcal{R}^d$, $a = (a^1, \ldots, a^d)$, $b = (b^1, \ldots, b^d)$. On conviendra d'écrire $a \leqslant b$ ($a < b$) si $a^k \leqslant b^k$ (resp. $a^k < b^k$) pour tous les $k = 1, \ldots, d$.

On appellera *intervalle d-dimensionnel* (*ou intervalle de* $\mathcal{R}^d$) l'ensemble $I\,[a, b[= \{x : a \leqslant x < b\}$. On définit de façon analogue les intervalles fermés $I\,[a, b] = \{x : a \leqslant x \leqslant b\}$ et les intervalles ouverts $I\,]a, b[= \{x : a < x < b\}$. On exprime sans peine la probabilité qu'un vecteur aléatoire ξ tombe dans un intervalle d-dimensionnel à l'aide de la fonction de répartition. A cet effet posons

$$\Delta^k_{[t,s[} f(x^1, x^2, \ldots, x^d) = f(x^1, x^2, \ldots, x^{k-1}, s, x^{k+1}, \ldots, x^d) - f(x^1, \ldots, x^{k-1}, t, x^{k+1}, \ldots, x^d).$$

De toute évidence

$$\Delta^k_{[t,s[} F(x) = \mathsf{P}\Big(\bigcap_{\substack{j=1\\ j\neq k}}^{d} \{\xi^j < x^j\} \cap \{t \leqslant \xi^k < s\}\Big),$$

et on s'assure sans peine que

$$F(I\,[a, b[) = \Delta^1_{[a^1, b^1[} \Delta^2_{[a^2, b^2[} \ldots \Delta^d_{[a^d, b^d[} F(x) = \mathsf{P}(\xi \in I\,[a, b[). \quad (1)$$

THÉORÈME 1. *La fonction de répartition $F(x)$ possède les propriétés suivantes:*

a) $0 \leqslant F(x) \leqslant 1$;

b) *si $x \leqslant y$, alors $F(x) \leqslant F(y)$*;

c) $\Delta^1_{[a^1, b^1[} \Delta^2_{[a^2, b^2[} \ldots \Delta^d_{[a^d, b^d[} F(x) \geqslant 0 \quad \forall a \leqslant b$;

d) *la fonction* $F(x)$ *est continue à gauche, c'est-à-dire* $F(x-0) = F(x)$;

e) $F(x) \to 0$ *si* $\min_k x^k \to -\infty$ *et* $F(x) \to 1$ *si* $\min_k x^k \to +\infty$.

Démonstration. a) est évident. Si $x \leqslant y$, alors $F(y) - F(x) = \mathsf{P}(\{\xi < y\} \setminus \{\xi < x\})$, d'où suit b). c) découle de (1). Soit $x_n \leqslant x_{n+1} \leqslant x$ et $\lim x_n = x$. Les événements $A_n = \{\xi \leqslant x\} \setminus \{\xi < x_n\}$ forment une suite monotone décroissante et $\bigcap_{n=1}^{\infty} A_n = \varnothing$. Donc (cf. (6), § 1) $F(x) - F(x_n) = \mathsf{P}(A_n) \to 0$ lorsque $n \to \infty$, ce qui démontre d). D'autre part, si $\min_k x_n^k \to -\infty$, il existe un j tel que $x_n^j \to -\infty$. Posons $z_n = \sup\{x_k^j,\ k = n, n+1, \ldots\}$, il vient $z_n \to -\infty$, z_n est monotone non croissante et $\bigcap_{n=1}^{\infty} \{\xi^j < z_n\} = \varnothing$. Donc $F(x_n) \leqslant \mathsf{P}\{\xi^j < z_n\} \to 0$. On démontre de façon analogue la deuxième partie de e) (à remarquer que $1 - F(x) = \mathsf{P}(\bigcup_{k=1}^{\infty} \{\xi^k \geqslant x^k\}))$. ■

Une fonction de $\mathcal{R}^d$ est *fonction de répartition* (*dans* $\mathcal{R}^d$) si elle possède les propriétés a) à e). Elle est *fonction monotone croissante dans* $\mathcal{R}^d$ si elle ne possède que les propriétés b) et c). Il convient de faire une distinction entre une fonction monotone croissante dans $\mathcal{R}^d$ et une fonction de d variables monotone relativement à chaque variable et ne possédant que la propriété *b*).

Faisons la remarque suivante sur les sauts éventuels de la fonction de répartition $F(x)$. Celle-ci est monotone comme fonction d'une seule variable fixée x^k, donc les limites de $F(x^1 + 0, x^2 + 0, \ldots, x^d + 0)$ existent en chaque point x.

Soit l'hyperplan $H_c^k = \{x : x^k = c\}$. Les hyperplans $H_{c_1}^k$ et $H_{c_2}^k$ ne se coupent pas pour $c_1 \neq c_2$. Il existe donc un ensemble au plus dénombrable de valeurs c_r^k, $r = 1, 2, \ldots$, telles que

$$F(+\infty, \ldots, +\infty, c+0, +\infty, \ldots, +\infty) - F(-\infty, \ldots, -\infty, c, -\infty, \ldots, -\infty) = 0$$

pour tous les $c \neq c_r^k$.

Appelons $H_{c_r^k}^k$ *hyperplans de discontinuité de la fonction* $F(x)$.

Donc si $x \notin H^0$, où $H^0 = \bigcup_{k=1}^{\infty} \bigcup_{r=1}^{\infty} H_{c_r^k}^k$, alors $F(x)$ est continue en x et si $a \notin H^0$ et $b^0 \notin H^0$, alors $F(I[a_n, b_n[) \to F(I[a, b[)$ lorsque $a_n \to a$ et $b_n \to b$.

Espace probabilisé finidimensionnel. Soit $\{\mathcal{R}^d, \mathfrak{S}, \mathsf{P}\}$ un espace probabilisé, $\mathfrak{S} \supset \mathfrak{B}^d$, où $\mathfrak{B}^d$ est la tribu des boréliens de $\mathcal{R}$.

Si l'on pose

$$F(x) = \mathbf{P}\{I] -\infty, x[\},$$

alors $F(x)$ sera fonction de répartition.

THÉORÈME 2. *Soit donnée dans $\mathcal{R}^d$ une fonction de répartition $F(x)$. On peut alors définir un espace probabilisé $\{\mathcal{R}^d, \mathfrak{S}, \mathbf{P}\}$ tel que $\mathfrak{S} \supset \mathfrak{B}^d$ et $\mathbf{P}\{I]-\infty, x[\} = F(x)$, la probabilité $\mathbf{P}$ étant définie de façon univoque sur $\mathfrak{B}^d$.*

La démonstration de cette proposition est basée sur des théorèmes généraux du prolongement des mesures. Enonçons-les.

DÉFINITIONS. 1. *Une classe non triviale d'ensembles $\mathfrak{M}$ est un semi-anneau si quel que soit $\Delta_i \in \mathfrak{M}$, $i = 1, 2$,*

$$\Delta_1 \cap \Delta_2 \in \mathfrak{M}, \quad \Delta_1 \setminus \Delta_2 = \bigcup_{k=1}^{r} \Delta^k, \text{ où } \Delta^k \in \mathfrak{M}.$$

2. *Une fonction d'ensembles additive non négative m définie sur un semi-anneau $\mathfrak{M}$ est une prémesure sur $\{\mathfrak{M}, m\}$.*

3. *Une fonction d'ensembles m' définie sur une classe d'ensembles $\mathfrak{M}'$ prolonge une fonction d'ensembles m définie sur $\mathfrak{M}$ si $\mathfrak{M}' \supset \mathfrak{M}$ et $m'(\Delta) = m(\Delta)$, $\Delta \in \mathfrak{M}$.*

4. *Une prémesure est σ-finie sur X si existe une suite Δ_k telle que $X = \bigcup_{k=1}^{\infty} \Delta_k$, $\Delta_k \in \mathfrak{M}$.*

THÉORÈME 3 (théorème de prolongement de la mesure). *Une prémesure m admet un prolongement $\{\mathfrak{S}, q\}$, où $\mathfrak{S}$ est une tribu et q une mesure sur $\mathfrak{S}$ si et seulement si elle est semi-additive, c'est-à-dire si $\bigcup_{k=1}^{\infty} \Delta_k \supset \Delta$ entraîne*

$$m(\Delta) \leqslant \sum_{k=1}^{\infty} m(\Delta_k) \tag{2}$$

quels que soient Δ et Δ_k de $\mathfrak{M}$. Si de plus m est σ-finie, alors q est définie de façon unique sur $\sigma\{\mathfrak{M}\}$.

THÉORÈME 4. *Toute classe J d'intervalles semi-ouverts $I[a, b[$. $a, b \in \mathcal{R}^d$, forme un semi-anneau et $F(I[a, b[)$ est une prémesure sur J.*

En effet, $I[a, b[\cap I[c, d[= I[t, s[$, où $t = (t^1, \ldots, t^d)$, $s = (s^1, \ldots, s^d)$, $t^k = \max(a^k, c^k)$, $s^k = \min(b^k, d^k)$ et $I[t, s[= \varnothing$, si la condition $t \leqslant s$ n'est pas réalisée. D'autre part $I[a, b[\setminus I[c, d[= \{x : x^j \in [a^j, b^j[\setminus [c^j, d^j[, j = 1, \ldots, d\}$, on reconnaît ici la somme d'au plus 2^d intervalles. Montrons maintenant que $F(I[a, b[)$ est une fonction additive sur J. Si l'on divise l'intervalle $I[a, b[$ en deux : $I[a, c_1[$ et $I[c_1, b[$, où $c_1 = (c^1, b^2, \ldots, b^d)$,

$a^1 < c^1 < b^1$, on a

$$F(I[a, b[) = \Delta^1_{[a^1, c^1[} \Delta^2_{[a^2, b^2[} \dots \Delta^d_{[a^d, b^d[} F(x) +$$
$$+ \Delta^1_{[c^1, b^1[} \Delta^2_{[c^2, b^2[} \dots \Delta^d_{[c^d, b^d[} F(x) = F(I[a, c_1[) + F(I[c_1, b[).$$

On obtiendrait la même égalité en divisant $I\ [a, b[$ en deux intervalles par une bipartition de l'un quelconque des côtés $[a^k, b^k[$. L'additivité de la fonction $F(I)$ se démontre par récurrence pour une partition quelconque de l'intervalle I. ■

Passons à la démonstration du théorème 2. Il reste à montrer que la fonction $F(I)$ possède la propriété de semi-additivité (2).

Soit $I = I[a_0, b_0[$, $I_n = I[a_n, b_n[$ et $I \subset \bigcup\limits_1^\infty I_n$. La fonction $F(x)$ étant continue à gauche, on peut trouver un $\varepsilon = (\varepsilon_k, \dots, \varepsilon_k)$, $\varepsilon_k > 0$, tel que

$$0 \leqslant F(I[a_k - \varepsilon_k, b_k[) - F(I[a_k, b_k[) < \frac{\eta}{2^k}, \quad \eta > 0 \ (k = 1, 2, \dots).$$

Les intervalles ouverts $I]a_k - \varepsilon_k, b_k[$ couvrent l'intervalle fermé $I[a_0, b_0 - \varepsilon]$. En vertu du théorème de Heine-Borel on peut en extraire un sous-recouvrement fini, par exemple $\{I\ ]a_k - \varepsilon_k, b_k[, k = 1, \dots, n\}$. La suite d'intervalles $I[a_k - \varepsilon_k, b_k[$, $k = 1, \dots, n$, couvre alors l'intervalle $I[a_0, b_0 - \varepsilon[$. Les ensembles disjoints

$$I[a_0, b_0 - \varepsilon[\cap \{I[a_k - \varepsilon_k, b_k[\setminus \bigcup_{i=1}^{k-1} I[a_i - \varepsilon_i, b_i[\}$$

$(k = 1, 2, \dots, n)$ sont sommes des intervalles semi-ouverts disjoints I_{kj} $(j = 1, \dots, d)$. Donc $I[a_0, b_0 - \varepsilon[= \bigcup\limits_{k=1}^{n} \bigcup\limits_{j=1}^{d_k} I_{kj}$ et

$$F(I[a_0, b_0 - \varepsilon[) = \sum_{k=1}^{n} \sum_{j=1}^{d_k} F(I_{kj}) \leqslant \sum_{k=1}^{n} F(I[a_k - \varepsilon_k, b_k[) \leqslant$$
$$\leqslant \sum_{k=1}^{\infty} F(I[a_k - \varepsilon_k, b_k[) \leqslant \sum_{k=1}^{\infty} F(I_k) + \eta.$$

En passant à la limite pour $\varepsilon \to 0$ on obtient

$$F(I[a_0, b_0[) \leqslant \sum_{k=1}^{\infty} F(I_k) + \eta$$

ou, puisque η est arbitraire,

$$F(I[a, b_0[) \leqslant \sum_{k=1}^{\infty} F(I_k). \quad ■$$

Théorème de Kolmogorov. Etudions le problème suivant. Soit donné un ensemble d'espaces probabilisables $\{X_s, \mathfrak{B}_s\}$, $s \in S$, et une collection de mesures probabilistes $\{m_{s_1, s_2, \ldots, s_n};\ n = 1, 2, \ldots;\ s_k \in S\}$, où $m_{s_1, s_2, \ldots, s_n}$ est une mesure sur le produit cartésien $\prod_{k=1}^{n} \{X_{s_k}, \mathfrak{B}_{s_k}\}$. On demande de construire un espace probabilisé $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ et une famille $\{\zeta_s, s \in S\}$, où ζ_s est un élément aléatoire de $\{X_s, \mathfrak{B}_s\}$, tels que toute suite $(\zeta_{s_1}, \zeta_{s_2}, \ldots, \zeta_{s_n})$ (n quelconque) possède une répartition donnée $m_{s_1, \ldots, s_n}$ sur $\prod_{k=1}^{n} \{X_{s_k}, \mathfrak{B}_{s_k}\}$.

Il est clair que la collection de mesures $m_{s_1, \ldots, s_n}$ ne peut être arbitraire. En effet, si le problème admet une solution et

$$m_{s_1, s_2, \ldots, s_n}\left(\prod_{k=1}^{n} B_{s_n}\right) = \mathsf{P}\left(\bigcap_{k=1}^{n} \{\zeta_{s_k} \in B_{s_k}\}\right), \tag{3}$$

il est évident que

$$m_{s_1, \ldots, s_{n+r}}\left(\prod_{k=1}^{n+r} B_{s_k}\right) = m_{s_1, \ldots, s_n}\left(\prod_{k=1}^{n} B_{s_k}\right) \tag{4}$$

si $B_{s_j} = X_{s_j}$, $j = n+1, \ldots, n+m$ et

$$m_{s_1, \ldots, s_n}\left(\prod_{k=1}^{n} B_{s_k}\right) = m_{s_{j_1}, \ldots, s_{j_n}}\left(\prod_{k=1}^{n} B_{s_{j_k}}\right), \tag{5}$$

où $(j_1, j_2, \ldots, j_n)$ est une permutation des nombres $(1, 2, \ldots, n)$.

Les relations (4) et (5) sont appelées *conditions de compatibilité* de la famille de répartitions $\{m_{s_1, \ldots, s_n}, n = 1, 2, \ldots, s_k \in S\}$.

Théorème 5 (théorème de Kolmogorov). *Soient X_s espaces polonais, $\mathfrak{B}_s$ la tribu des boréliens de X_s. Pour toute famille compatible de répartitions $\{m_{s_1, \ldots, s_n}, n = 1, 2, \ldots, s_k \in S\}$ on peut construire un espace probabilisé $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ et une famille de variables aléatoires $\{\zeta_s, s \in S\}$, tels que $m_{s_1, \ldots, s_n}$ soit répartition de la suite $\{\zeta_{s_1}, \ldots, \zeta_{s_n}\}$.*

Avant d'entamer la démonstration du théorème voici quelques remarques et constructions qui nous seront utiles pour la suite.

Soit Ω l'espace des fonctions $\omega = \omega(s)$, $s \in S$, à valeurs dans X_s $(\Omega = \prod_{s \in S} X_s)$.

Définition. *Tout ensemble $C_{s_1, \ldots, s_n}(B^{(s_1, \ldots, s_n)})$ de la forme*

$$C_{s_1, \ldots, s_n}(B^{(s_1, \ldots, s_n)}) = \{\omega : (\omega(s_1), \ldots, \omega(s_n)) \in B^{(s_1, \ldots, s_n)}\},$$

où $B^{(s_1, \ldots, s_n)} \in \sigma\{\mathfrak{B}_{s_k},\ k=1, 2, \ldots, n\}$, *est appelé cylindrique ou, plus exactement, ensemble cylindrique dans* Ω *à base* $B^{(s_1, \ldots, s_n)}$ *sur les coordonnées* $s_1, \ldots, s_n$.

A $s_1, s_2, \ldots, s_n$ fixes les ensembles cylindriques $C_{s_1, \ldots, s_n} \times \times (B^{(s_1, \ldots, s_n)})$ et les ensembles de $\sigma\{\mathfrak{B}_{s_k},\ k=1, \ldots, n\}$ sont isomorphes: tout ensemble $B \in \sigma\{\mathfrak{B}_{s_k},\ k=1, \ldots, n\}$ est base d'un ensemble cylindrique $C_{s_1, \ldots, s_n}(B)$; à des bases différentes sont associés des ensembles cylindriques différents; à la somme, la différence ou l'intersection des bases correspondent une somme, une différence ou une intersection d'ensembles cylindriques. Ceci découle immédiatement de la définition de l'ensemble cylindrique.

A noter qu'un même ensemble cylindrique est susceptible d'être défini sur des collections différentes de coordonnées.

Ainsi il est évident que

$$C_{s_1, \ldots, s_n}(B) = C_{s_1, \ldots, s_n, \ldots, s_{n+r}}(B \times X_{n+1} \times \ldots \times X_{s_{n+r}}).$$

On voit sans peine que deux ensembles cylindriques quelconques $C_{s_1, \ldots, s_n}(B)$ et $C_{s'_1, \ldots, s'_r}(B)$ peuvent toujours être considérés comme des ensembles cylindriques sur une même suite de coordonnées $s''_1, \ldots, s''_r$ contenant aussi bien $s_1, \ldots, s_n$ que $s'_1, \ldots, s'_p$. D'où il suit qu'en effectuant des opérations algébriques sur un nombre fini d'ensembles cylindriques on peut admettre qu'ils sont donnés sur une suite fixe de coordonnées. Donc la classe des ensembles cylindriques engendre une algèbre d'ensembles.

Ajoutons à cela que si S est un ensemble infini et les X_s contiennent au moins deux points, la classe des ensembles cylindriques n'est pas une tribu.

En effet, l'ensemble $\bigcup_1^\infty C_{s_n}(\{x_{s_n}\})$, où x_{s_n} est un ensemble à un seul point, $x_{s_n} \in X_{s_n}$, n'est pas cylindrique.

Soit X_k un ensemble métrique complet, ρ_k la métrique correspondante, $k=1, 2, \ldots, n$. Munissons l'espace $Y = \prod_{k=1}^n X_k$ de la métrique

$$\rho(y_1, y_2) = \sqrt{\sum_{k=1}^n \rho_k^2(x_1^k, x_2^k)},$$

où $y_i = (x_i^1, \ldots, x_i^n)$, $x_i^k \in X_k$.

Les points x_i^k $(k = 1, \ldots, n)$ seront appelés *coordonnées* du point $y_i \in Y$. Il est aisé de voir qu'une suite de points y_n est convergente

si et seulement si les coordonnées des points y_n le sont dans les espaces respectifs. D'où il suit que Y est un espace métrique complet. Il est séparable, puisque tout ensemble dénombrable de points de la forme $(x^1_{r_1}\, x^2_{r_2}, \ldots, x^n_{r_n}), x^k_{r_k} \in Z_k$, où Z_k est un ensemble dénombrable partout dense dans X_k, forme un réseau partout dense dans Y.

Pour démontrer le théorème de Kolmogorov on aura besoin du

THÉORÈME 6. *Si X est un espace polonais, m une mesure probabiliste sur la tribu des boréliens de X, alors pour tout $\varepsilon > 0$ et $B \in \mathfrak{B}$ il existe un compact K, $K \in \mathfrak{B}$, tel que $m(B \setminus K) < \varepsilon$.*

D é m o n s t r a t i o n d u t h é o r è m e d e K o l m o g o r o v. Considérons de nouveau l'espace Ω des fonctions $\omega = \omega(s)$, $\omega(s) \in X_s$, $s \in S$, et posons $\mathsf{P}'(C) = m_{s_1, \ldots, s_n}(B^{(s_1, \ldots, s_n)})$ pour un ensemble cylindrique quelconque $C = C_{s_1, \ldots, s_n}(B^{(s_1, \ldots, s_n)})$. Des conditions de compatibilité des mesures il suit que $\mathsf{P}'(C)$ est définie de façon unique. Soit C_k, $k = 1, \ldots, n, \ldots$, une suite d'ensembles cylindriques. Sans restreindre la généralité on peut admettre qu'ils sont donnés par les bases $B_k^{(s_1, \ldots, s_p)}$ sur la même suite de coordonnées $(s_1, \ldots, s_p, \ldots)$. Les ensembles cylindriques sur une suite fixe de coordonnées et leurs bases étant isomorphes, on constate que $\mathsf{P}'(C)$ est une mesure finiment additive. Montrons qu'elle est prolongeable à une mesure P sur $\{\Omega, \mathfrak{S}\}$, où $\mathfrak{S}$ est une tribu. En vertu du théorème du prolongement de la mesure il suffit simplement de vérifier que P' est semi-additive, c'est-à-dire, si $C_0 \subset \bigcup_{n=1}^{\infty} C_n$, que

$$\mathsf{P}'(C_0) \leqslant \sum_{k=1}^{\infty} \mathsf{P}'(C_n).$$

Soit $C_0 = \bigcup_{n=1}^{\infty} C_n$ $(C_k \cap C_r = \varnothing$ pour $k \neq r)$. Montrons que

$$\mathsf{P}'(C_0) = \sum_{n=1}^{\infty} \mathsf{P}'(C_n). \tag{6}$$

D'où résultera la proposition annoncée. Posons $C'_n = C_0 \setminus \bigcup_1^n C_k$. Les ensembles C'_n forment une suite monotone non décroissante d'ensembles cylindriques disjoints. Comme

$$\mathsf{P}'(C_0) = \sum_{k=1}^{n} \mathsf{P}'(C_k) + \mathsf{P}'(C'_n),$$

pour prouver (6) il suffit de montrer que $\lim \mathsf{P}'(C'_n) = 0$.

On raisonnera par l'absurde, c'est-à-dire on supposera que $\lim \mathsf{P}'(C'_n) = a > 0$. Soit B_n la base de l'ensemble cylindrique C'_n et supposons que C'_n est situé sur les coordonnées $s_1, s_2, \ldots, s_{r_n}$. Sans

nuire à la généralité on peut admettre que lorsque n croît, la collection des points correspondants $(s_1, \ldots, s_{r_n})$ ne décroît pas. En vertu du théorème 6 il existe un compact $K_n \subset B_n$ tel que

$$m_{s_1, \ldots, s_{r_n}}(B_n \setminus K_n) < \frac{a}{2^{n+1}}, \quad n = 1, 2, \ldots$$

Soient Q_n ensemble cylindrique sur les coordonnées $s_1, \ldots, s_{r_n}$ de base K_n, $F_n = \bigcap_{r=1}^{n} Q_r$ et D_n base de l'ensemble F_n. De toute évidence D_n est un compact dans $\prod_{k=1}^{r_n} X_{s_k}$ puisque D_n est l'intersection d'ensembles fermés dont au moins un, pour fixer les idées K_n, est compact.

Les ensembles F_n étant monotones décroissants, de l'appartenance $\omega(s) \in F_{n'}$, $n' > n$, il résulte que $\omega(s) \in F_n$. Donc si $(x_{s_1}, x_{s_2}, \ldots, x_{s_{r_n}}, \ldots, x_{s_{r_{n+p}}}) \in D_{n+p}$, alors $(x_{s_1}, x_{s_2}, \ldots, x_{s_{r_n}}) \in \in D_n$.

Les ensembles F_n sont visiblement non vides. Bien plus, puisque

$$C'_n \setminus F_n = \bigcup_{j=1}^{n} (C'_n \setminus Q_j) \subset \bigcup_{j=1}^{n} (C'_j \setminus Q_j),$$

alors

$$\mathsf{P}'(C'_n \setminus F_n) < \sum_{1}^{n} \mathsf{P}(C'_j \setminus Q_j) = \sum_{1}^{n} m_{s_1 \ldots s_{r_j}}(B_j) \setminus K_j) \leqslant \frac{a}{2},$$

d'où il vient

$$\lim \mathsf{P}'(F_n) = \lim \mathsf{P}'(C'_n) - \lim \mathsf{P}'(C'_n \setminus F_n) \geqslant \frac{a}{2}.$$

Repérons un point quelconque $(x_1^n, \ldots, x_{r_n}^n)$ dans chaque ensemble D_n. Quel que soit k, la suite de points x_k^n, $n = 1, 2, \ldots$, appartient à un compact dans X_{s_k} et la suite $(x_1^{n+j}, \ldots, x_{r_n}^{n+j})$, $j = 0, 1, 2, \ldots$ est contenue dans D_n. Utilisons le processus diagonal pour trouver une suite d'indices n_j tels que pour tout k la suite $x_k^{n_j}$ converge vers une limite x_k^0. L'ensemble D_n étant fermé, il suit que $(x_1^0, x_2^0, \ldots \ldots, x_{r_n}^0) \in D_n$ quel que soit n.

Soit la fonction $\omega(s)$ telle que $\omega(s_k) = x_k^0$, $k = 1, 2, \ldots$, et à valeurs quelconques dans les autres points. Alors pour tout n on a $\omega(s) \in F_n \subset C'_n$. Donc $\bigcap_{n=1}^{\infty} C'_n$ n'est pas vide, ce qui est contraire à l'hypothèse, d'où il suit que $\lim \mathsf{P}'(C'_n) = 0$ et la prémesure P' est prolongeable à une mesure P sur la tribu $\mathfrak{S}$ de tous les ensembles cylindriques de l'espace. On a ainsi construit l'espace probabilisé

$\{\Omega, \mathfrak{S}, \mathsf{P}\}$. Posons maintenant $\zeta_s = g_s(\omega) = \omega(s)$. Il vient

$$\mathsf{P}\{(\zeta_{s_1}, \ldots, \zeta_{s_n}) \in B^{(s_1, \ldots, s_n)}\} =$$
$$= \mathsf{P}\{(\omega(s_1), \ldots, \omega(s_n)) \in B^{(s_1, \ldots, s_n)}\} =$$
$$= \mathsf{P}'(C_{s_1, \ldots, s_n}(B^{(s_1, \ldots, s_n)})) = m_{s_1, \ldots, s_n}(B^{(s_1, \ldots, s_n)}). \blacksquare$$

§ 3. Probabilités conditionnelles

La formule élémentaire de la probabilité d'un événement A sachant B

$$\mathsf{P}(A \mid B) = \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B)}, \quad \mathsf{P}(B) > 0, \tag{1}$$

montre comment il convient de calculer la probabilité si la classe de possibles ou les conditions C sous lesquelles sont réalisées les expériences se modifient. Plus exactement, aux conditions C on ajoute la suivante: on n'envisage que les expériences dans lesquelles B est obligatoirement réalisé.

Posons le problème sous un angle plus large. Supposons qu'est effectuée une expérience E. Comment faut-il déterminer la probabilité des événements observables dans d'autres expériences si l'on admet que l'issue de E est fixée?

Avant de donner une définition formelle traitons quelques cas particuliers de ce problème. Comme l'expérience E est entièrement décrite par une tribu $\mathfrak{F}$ d'événements observables de E, on désignera la probabilité cherchée par $\mathsf{P}\{A \mid \mathfrak{F}\}$ et on l'appellera probabilité conditionnelle de l'événement A par rapport à la tribu $\mathfrak{F}$. Si $A \in \mathfrak{F}$, on posera tout naturellement $\mathsf{P}\{A \mid \mathfrak{F}\} = 1$ ou 0 selon que l'événement A s'est réalisé ou non.

Soient maintenant $\mathfrak{F}$ tribu simple, E_n ses atomes et $\mathsf{P}(E_n) > 0$. Si A est un événement sur $\mathfrak{S}$, on posera

$$\mathsf{P}\{A \mid \mathfrak{F}\} = \sum_n \mathsf{P}(A \mid E_n)\, \chi(E_n), \tag{2}$$

où $\mathsf{P}(A \mid E_n)$ est défini par la formule (1).

A noter que le trait singulier de cette définition est que la probabilité conditionnelle est une variable aléatoire dépendant du résultat de l'expérience réalisée. Ceci signifie formellement que la probabilité conditionnelle est une variable aléatoire $\mathfrak{F}$-mesurable.

Introduisons la notion d'espérance mathématique conditionnelle $\mathsf{E}\{\xi \mid \mathfrak{F}\}$ d'une variable aléatoire ξ. Il est naturel de poser

$$\mathsf{E}\{\xi \mid \mathfrak{F}\} = \int_\Omega \xi \mathsf{P}(d\omega \mid \mathfrak{F})$$

ou

$$\mathsf{E}\{\xi \mid \mathfrak{F}\} = \sum_n \frac{\chi(E_n)}{\mathsf{P}(E_n)} \int_{E_n} \xi\, d\mathsf{P}. \tag{3}$$

On supposera en outre que $\mathsf{E}\xi$ est fini. Si dans la dernière formule on pose $\xi = \chi(A)$, on obtient la probabilité conditionnelle de l'événement A :

$$\mathsf{E}\{\chi(A) \mid \mathfrak{F}\} = \sum_n \chi(E_n) \frac{\mathsf{P}(E_n \cap A)}{\mathsf{P}(E_n)} = \mathsf{P}\{A \mid \mathfrak{F}\},$$

de sorte que la probabilité conditionnelle est un cas particulier de l'espérance mathématique conditionnelle.

Soit η variable aléatoire $\mathfrak{F}$-mesurable bornée. On a alors $\eta = \sum_n a_n \chi(E_n)$. En multipliant (3) par η on obtient

$$\mathsf{E}(\eta \mathsf{E}\{\xi \mid \mathfrak{F}\}) = \sum_n a_n \int_{E_n} \xi\, d\mathsf{P} = \sum_n \int_{E_n} \eta\xi\, d\mathsf{P}$$

ou

$$\mathsf{E}(\eta \mathsf{E}\{\xi \mid \mathfrak{F}\}) = \mathsf{E}\eta\xi. \tag{4}$$

La dernière relation définit complètement l'espérance mathématique conditionnelle de la variable aléatoire ξ. En effet, supposons que $\check{\xi}$ est une variable $\mathfrak{F}$-mesurable telle que pour toute variable aléatoire $\mathfrak{F}$-mesurable bornée η on ait $\mathsf{E}(\eta\check{\xi}) = \mathsf{E}(\eta\xi)$. Comme $\check{\xi}$ est constante sur E_n, il vient $\check{\xi} = c_n$ pour $\omega \in E_n$. Soit $\eta = \chi(E_n)$. On a

$$\mathsf{E}(\eta\check{\xi}) = c_n \mathsf{E}\chi(E_n) = c_n \mathsf{P}(E_n) = \int_{E_n} \xi\, d\mathsf{P},$$

c'est-à-dire $\check{\xi}$ est confondue avec le second membre de (3).

Utilisons (4) pour généraliser la définition de l'espérance mathématique conditionnelle. Si dans (4) on pose $\eta = \chi(F)$, on obtient

$$\int_F \mathsf{E}\{\xi \mid \mathfrak{F}\}\, d\mathsf{P} = \int_F \xi\, d\mathsf{P}. \tag{5}$$

DÉFINITION. *On appelle espérance mathématique conditionnelle* $\mathsf{E}\{\xi \mid \mathfrak{F}\}$ *d'une variable aléatoire* ξ (*on admet que* $\mathsf{E}\xi$ *existe*), *par rapport à une tribu* $\mathfrak{F}$, *une variable aléatoire* $\mathfrak{F}$*-mesurable vérifiant* (5) *pour tout* $F \in \mathfrak{F}$.

THÉORÈME 1. *L'espérance mathématique conditionnelle d'une variable aléatoire quelconque* ($\mathsf{E}\xi$ *est définie*) *existe et est unique* (mod P).

Démonstration. Le second membre de la formule (5) est une charge $\varphi(F)$ σ-finie sur $\mathfrak{F}$, absolument continue par rapport à la mesure P. D'après le théorème de Radon-Nikodym il existe

une fonction $g(u)$ $\mathfrak{F}$-mesurable telle que

$$\varphi(F) = \int_F g(u)\,\mathsf{P}(du).$$

Reste à poser $\mathsf{E}\{\xi \mid \mathfrak{F}\} = g(u)$. Prouvons l'unicité (mod P) de l'espérance mathématique conditionnelle. Si existent deux variables aléatoires $\check{\xi}_1$ et $\check{\xi}_2$ vérifiant la définition de l'espérance mathématique conditionnelle, pour tout $F \in \mathfrak{F}$ on a

$$\int_F (\check{\xi}_1 - \check{\xi}_2)\,d\mathsf{P} = 0,$$

ce qui en vertu de la $\mathfrak{F}$-mesurabilité de $\check{\xi}_1 - \check{\xi}_2$ n'est possible que pour $\check{\xi}_1 = \check{\xi}_2$ (mod P). ■

DÉFINITION. *On appelle probabilité conditionnelle d'un événement A par rapport à une tribu $\mathfrak{F}$ la variable aléatoire*

$$\mathsf{P}\{A \mid \mathfrak{F}\} = \mathsf{E}\{\chi(A) \mid \mathfrak{F}\}.$$

$\mathsf{P}\{A \mid \mathfrak{F}\}$ se définit directement de la manière suivante.

La probabilité conditionnelle $\mathsf{P}\{A \mid \mathfrak{F}\}$ est une variable aléatoire $\mathfrak{F}$-mesurable vérifiant

$$\int_F \mathsf{P}\{A \mid \mathfrak{F}\}\,d\mathsf{P} = \mathsf{P}(A \cap F) \tag{6}$$

pour tout $F \in \mathfrak{F}$.

Le théorème nous apprend que $\mathsf{P}\{A \mid \mathfrak{F}\}$ existe et est définie de façon unique (mod P) pour tout A ($A \in \mathfrak{S}$).

Il est évident que (4) a lieu pour toute variable aléatoire $\mathfrak{F}$-mesurable dès que $\eta\xi$ est intégrable. La relation (4) admet l'interprétation suivante. Supposons que $\mathsf{E}\xi^2 < \infty$. On traitera les variables aléatoires comme des vecteurs dans un espace hilbertien $\mathcal{L}_2(\Omega, \mathfrak{S}, \mathsf{P})$. Posons pour abréger $\check{\xi} = \mathsf{E}\{\xi \mid \mathfrak{F}\}$ et soit H le sous-espace $\mathcal{L}_2(\Omega, \mathfrak{F}, \mathsf{P})$ de toutes les variables aléatoires $\mathfrak{F}$-mesurables à moments finis du second ordre. H est un sous-espace vectoriel fermé de $\mathcal{L}_2(\Omega, \mathfrak{S}, \mathsf{P})$, $\check{\xi} \in H$, et pour tout $\eta \in H$ on a $\mathsf{E}\eta(\check{\xi} - \xi) = 0$. Ce qui traduit l'orthogonalité du vecteur $\check{\xi} - \xi$ au sous-espace H, c'est-à-dire $\check{\xi}$ est la projection du vecteur ξ sur H. Donc, si $\mathsf{E}\xi^2 < \infty$, l'espérance mathématique conditionnelle $\mathsf{E}\{\xi \mid \mathfrak{F}\}$ est la projection de ξ sur H.

Enumérons quelques propriétés des espérances mathématiques conditionnelles. Nous conviendrons que les variables aléatoires envisagées plus bas possèdent des espérances mathématiques finies ou infinies et que toutes les égalités ont presque sûrement lieu.

a) *Si une variable aléatoire ξ est $\mathfrak{F}$-mesurable, on a*

$$\mathsf{E}\{\xi \mid \mathfrak{F}\} = \xi.$$

En effet, (5) est triviale si l'on pose $\mathsf{E}\{\xi \mid \mathfrak{F}\} = \xi$. En particulier, si $A \in \mathfrak{F}$, il vient

$$\mathsf{P}\{A \mid \mathfrak{F}\} = \chi(A). \tag{7}$$

b) *Si* ξ_1, ξ_2 *prennent des valeurs de même signe ou possèdent une espérance mathématique finie, alors*

$$\mathsf{E}\{\xi_1 + \xi_2 \mid \mathfrak{F}\} = \mathsf{E}\{\xi_1 \mid \mathfrak{F}\} + \mathsf{E}\{\xi_2 \mid \mathfrak{F}\}.$$

Posons $\check{\xi}_i = \mathsf{E}\{\xi_i \mid \mathfrak{F}\}$. Pour tout $F \in \mathfrak{F}$

$$\int_F (\check{\xi}_1 + \check{\xi}_2)\, d\mathsf{P} = \int_F \check{\xi}_1\, d\mathsf{P} + \int_F \check{\xi}_2\, d\mathsf{P} = \int_F \xi_1\, d\mathsf{P} + \int_F \xi_2\, d\mathsf{P} = \int_F (\xi_1 + \xi_2)\, d\mathsf{P}.$$

L'assertion proposée découle de l'unicité (mod P) de l'espérance mathématique conditionnelle.

COROLLAIRE. *Si* $A \cap B = \varnothing$, *alors*

$$\mathsf{P}\{A \cup B \mid \mathfrak{F}\} = \mathsf{P}\{A \mid \mathfrak{F}\} + \mathsf{P}\{B \mid \mathfrak{F}\} \ (\mathrm{mod}\ \mathsf{P}).$$

c) *Si* $\xi \leqslant \eta$, *alors*

$$\mathsf{E}\{\xi \mid \mathfrak{F}\} \leqslant \mathsf{E}\{\eta \mid \mathfrak{F}\}$$

sous réserve que l'un au moins des membres de cette inégalité ait un sens.

Si $\eta \geqslant 0$, alors

$$\int_F \mathsf{E}\{\eta \mid \mathfrak{F}\}\, d\mathsf{P} \geqslant 0$$

pour tout $F \in \mathfrak{F}$. Donc

$$\mathsf{E}\{\eta \mid \mathfrak{F}\} \geqslant 0.$$

Généralisons. En posant $\eta = \xi + (\eta - \xi)$ et en utilisant b) on obtient

$$\mathsf{E}\{\eta \mid \mathfrak{F}\} = \mathsf{E}\{\xi \mid \mathfrak{F}\} + \mathsf{E}\{\eta - \xi \mid \mathfrak{F}\} \geqslant \mathsf{E}\{\xi \mid \mathfrak{F}\}.$$

d) *Si* ξ_n, $n = 1, 2, \ldots$, *est une suite monotone non décroissante de variables aléatoires non négatives, alors*

$$\lim \mathsf{E}\{\xi_n \mid \mathfrak{F}\} = \mathsf{E}\{\lim \xi_n \mid \mathfrak{F}\}.$$

En effet, d'après le théorème de l'intégration des suites monotones

$$\int_F \lim \mathsf{E}\{\xi_k \mid \mathfrak{F}\}\, d\mathsf{P} = \lim \int_F \mathsf{E}\{\xi_n \mid \mathfrak{F}\}\, d\mathsf{P} \leqslant \lim \int_F \xi_n\, d\mathsf{P} = \int_F \lim \xi_n\, d\mathsf{P}.$$

COROLLAIRE. *Si* A_n, $n = 1, 2, \ldots$, *sont deux à deux disjoints, alors*

$$\mathsf{P}\{\bigcup_1^{\infty} A_n \mid \mathfrak{F}\} = \sum_{n=1}^{\infty} \mathsf{P}\{A_n \mid \mathfrak{F}\}. \tag{8}$$

Remarque importante. Les probabilités conditionnelles $\mathsf{P}\{A \mid \mathfrak{F}\} = \mathsf{P}_{\mathfrak{F}}\{A, \omega\}$ sont fonctions de deux arguments, A et ω ($A \in \mathfrak{S}$, $\omega \in \Omega$). L'égalité (8) a presque sûrement lieu et de plus l'ensemble exclusif N des ω tels que (8) ne soit pas réalisée dépend de la suite A_n. Donc, en général, on ne peut affirmer que $\mathsf{P}_{\mathfrak{F}}(A, \omega)$ est une mesure pour des ω.

e) *Si les espérances mathématiques de* ξ *et* $\alpha\xi$ *ont un sens,* α *étant une variable aléatoire* $\mathfrak{F}$*-mesurable, alors*

$$\mathsf{E}\{\alpha\xi \mid \mathfrak{F}\} = \alpha\mathsf{E}\{\xi \mid \mathfrak{F}\}. \tag{9}$$

En vertu de b) on peut faire la démonstration en se limitant à l'hypothèse que > 0 et $\alpha \geqslant 0$.

La relation (9) a été établie précédemment pour α et ξ discrets dans une forme différente mais équivalente (cf. formule (4)). Construisons dans le cas général une suite monotone croissante de variables aléatoires discrètes non négatives ξ_n convergentes vers ξ pour chaque ω et une suite monotone croissante de variables aléatoires discrètes non négatives $\mathfrak{F}$-mesurables α_n convergentes vers α. Faisons $\alpha = \alpha_n$, $\xi = \xi_m$ dans (9) et passons à la limite d'abord pour $m \to \infty$ et ensuite pour $n \to \infty$. En se servant de d) on obtient la proposition annoncée. ■

Corollaire. *Si* $F \in \mathfrak{F}$, *alors*

$$\mathsf{P}\{A \cap F \mid \mathfrak{F}\} = \chi(F)\,\mathsf{P}\{A \mid \mathfrak{F}\}. \tag{10}$$

L'espérance mathématique conditionnelle itérée possède une propriété d'« absorption » qui est importante et souvent utilisée.

Théorème 2. *Si* $\mathfrak{F}_1 \subset \mathfrak{F}_2$, *alors*

$$\mathsf{E}\{\mathsf{E}\{\xi \mid \mathfrak{F}_2\} \mid \mathfrak{F}_1\} = \mathsf{E}\{\xi \mid \mathfrak{F}_1\}.$$

Démonstration. $F \in \mathfrak{F}_1$ entraîne $F \in \mathfrak{F}_2$, donc

$$\int_F \mathsf{E}\{\mathsf{E}\{\xi \mid \mathfrak{F}_2\} \mid \mathfrak{F}_1\}\, d\mathsf{P} = \int_F \mathsf{E}\{\xi \mid \mathfrak{F}_2\}\, d\mathsf{P} = \int_F \xi\, d\mathsf{P} = \int_F \mathsf{E}\{\xi \mid \mathfrak{F}_1\}\, d\mathsf{P}.$$

Une comparaison des membres extrêmes des égalités nous donne le résultat annoncé. On remarquera que l'égalité $\mathsf{E}\{\mathsf{E}\{\xi \mid \mathfrak{F}_1\} \mid \mathfrak{F}_2\} = \mathsf{E}\{\xi \mid \mathfrak{F}_1\}$ ($\mathfrak{F}_1 \subset \mathfrak{F}_2$) est triviale. En effet, la quantité $\mathsf{E}\{\xi \mid \mathfrak{F}_1\}$ est $\mathfrak{F}_1$- et *a fortiori* $\mathfrak{F}_2$-mesurable. Cette égalité découle alors de a). ■

Soit maintenant une expérience décrite par un élément aléatoire ζ, $\zeta = g(\omega)$, à valeurs dans $\{X, \mathfrak{B}\}$. L'espérance mathématique conditionnelle $\mathsf{E}\{\xi \mid \zeta\}$ de ξ par rapport à ζ est la valeur moyenne qu'elle prend pour ζ fixe. D'où la

Définition. $\mathsf{E}\{\xi \mid \zeta\} = \mathsf{E}\{\xi \mid \mathfrak{F}_\zeta\}$ *où* $\mathfrak{F}_\zeta$ *est la tribu engendrée par l'élément aléatoire* ζ.

En s'inspirant de la définition initiale de l'espérance mathématique conditionnelle on peut énoncer la précédente sous la forme équivalente suivante :

$\mathsf{E}\{\xi \mid \zeta\}$ *est une variable aléatoire* $\mathfrak{F}_\zeta$*-mesurable vérifiant pour tout* $B \in \mathfrak{B}$ *la relation*

$$\int_{g^{-1}(B)} \mathsf{E}\{\xi|\zeta\}\, d\mathsf{P} = \int_{g^{-1}(B)} \xi\, d\mathsf{P}. \tag{11}$$

Théorème 3. *L'espérance mathématique conditionnelle* $\mathsf{E}\{\xi \mid \zeta\}$ *est une fonction* $\mathfrak{B}$*-mesurable de* ζ ; *c'est-à-dire il existe une fonction réelle* $\mathfrak{B}$*-mesurable* $h(x)$, $x \in X$, *telle que* $\mathsf{E}\{\xi \mid \zeta\} = h(\zeta)$ *et quel que soit* $B \in \mathfrak{B}$

$$\int_B h(x)\, \mathsf{P}g^{-1}(dx) = \int_{g^{-1}(B)} \xi\, d\mathsf{P}.$$

La première partie du théorème découle du théorème 5, § 1, la seconde de la règle de changement des variables (théorème 12, § 1). ∎

Signalons les propriétés suivantes des espérances mathématiques conditionnelles par rapport à des variables aléatoires, découlant du théorème précédent.

f) *Si* $\xi = h(\zeta)$, *où* $h(x)$ *est une fonction* $\mathfrak{B}$*-mesurable, alors* $\mathsf{E}\{\xi \mid \zeta\} = \xi$.

g) *Si* ζ_i *sont des éléments aléatoires dans* $\{X_i, \mathfrak{B}_i\}$, $i = 1, 2$, (ζ_1, ζ_2) *leur produit cartésien, alors*

$$\mathsf{E}\{\mathsf{E}\{\xi \mid (\zeta_1, \zeta_2)\} \mid \zeta_1\} = \mathsf{E}\{\xi \mid \zeta_1\}.$$

La proposition f) découle de e), la proposition g) du théorème 2.

Probabilités conditionnelles régulières. Posons $\mathsf{P}\{A \mid \mathfrak{F}\} = \mathsf{P}^{\mathfrak{F}}(A, \omega)$. Pour tout $A \in \mathfrak{S}$ la probabilité conditionnelle $\mathsf{P}^{\mathfrak{F}}(A, \omega)$ est définie de façon unique mais seulement avec la probabilité 1.

Définition. *On appelle probabilité conditionnelle régulière une fonction (si elle existe)* $p(A, \omega)$, $A \in \mathfrak{S}$, $\omega \in \Omega$, *telle que*

a) $p(A, \omega)$ *comme fonction de l'ensemble* A *est mesure probabiliste pour presque tous les* ω :

b) $p(A, \omega)$ *est* $\mathfrak{F}$*-mesurable pour* A *fixe et*

$$p(A, \omega) = \mathsf{P}^{\mathfrak{F}}(A, \omega) \pmod{\mathsf{P}}.$$

On pourrait donner des exemples où les probabilités conditionnelles régulières n'existent pas. Mais si celles-ci existent, les espérances mathématiques conditionnelles s'expriment en fonction d'elles par intégration.

THÉORÈME 4. *Si $p(A, \omega) = \mathsf{P}^{\mathfrak{F}}(A, \omega)$ est une probabilité conditionnelle régulière, alors*

$$\mathsf{E}\{\xi|\mathfrak{F}\} = \int f(\omega')\,\mathsf{P}(d\omega', \omega). \tag{12}$$

Démonstration. La classe des variables aléatoires ξ justiciables de cette formule est linéaire, fermée pour le passage à la limite dans les suites monotones croissantes de variables aléatoires non négatives (d'après le théorème de l'intégration des suites monotones) et contient les indicateurs $\chi(A)$, $A \in \mathfrak{S}$, d'où suit la formule (12) pour tous les ξ dont $\mathsf{E}\xi$ est finie. ■

Les répartitions conditionnelles régulières n'existant pas toujours, nous allons modifier cette notion dans une mesure qui permette l'abord de nombreux problèmes d'application.

Soient $\{X, \mathfrak{B}\}$ espace probabilisable, ζ élément aléatoire dans $\{X, \mathfrak{B}\}$, $\mathfrak{F}$ tribu, $\mathfrak{F} \subset \mathfrak{S}$.

DÉFINITION. *On appelle répartition conditionnelle régulière de l'élément aléatoire ζ par rapport à la tribu $\mathfrak{F}$ une fonction $Q(B, \omega)$, si elle existe, définie sur $\mathfrak{B} \times \Omega$ et telle que*

a) $Q(B, \omega)$ *est $\mathfrak{F}$-mesurable pour $B \in \mathfrak{B}$ fixe,*

b) $Q(B, \omega)$ *est presque sûrement une mesure probabiliste sur $\mathfrak{B}$ pour ω fixe,*

c) $Q(B, \omega) = \mathsf{P}\{(\zeta \in \mathfrak{B}) \mid \mathfrak{F}\}$ (mod P) *pour tout $B \in \mathfrak{B}$.*

La condition c) revient à exiger que pour tout $F \in \mathfrak{F}$

$$\int_F Q(B, \omega)\,\mathsf{P}(d\omega) = \mathsf{P}(\{\xi \in B\} \cap F).$$

THÉORÈME 5. *Soient X espace polonais, $\mathfrak{B}$ la tribu des boréliens de X, ζ élément aléatoire de $\{X, \mathfrak{B}\}$. Alors ζ admet une répartition conditionnelle régulière par rapport à une tribu $\mathfrak{F}$ ($\mathfrak{F} \subset \mathfrak{S}$).*

Démonstration. Soit q répartition de l'élément aléatoire ζ. On peut construire une suite monotone croissante de compacts K_n dans X telle que (cf. théorème 6, § 2) $q(X \setminus K_n) < \varepsilon_n$ et $\varepsilon_n \to 0$. Désignons par $\mathscr{C}(X)$ l'espace de toutes les fonctions continues bornées définies sur l'espace métrique X et munissons $\mathscr{C}(X)$ d'une métrique $\rho(f, g)$ en posant $r(f, g) = \|f - g\|$, $\|f\| = \sup_{x \in X} |f(x)|$. L'espace $\mathscr{C}(K_n)$ est séparable. Soit $\{f_{nk}(x),\ k = 1, 2, \ldots\}$ réseau dénombrable partout dense dans $\mathscr{C}(K_n)$. Prolongeons $f_{nk}(x)$ sur X de telle sorte que $\sup_{x \in K_n} |f_{nk}(x)| = \sup_{x \in X} |f_{nk}(x)|$.

Posons $\chi_n = \chi_n(\zeta)$, $\chi_n(x) = \chi(K_n, x)$. Des propriétés des espérances mathématiques conditionnelles il résulte que l'on peut exhiber un D_0 tel que $\mathsf{P}(D_0) = 0$ et pour $\omega \notin D_0$ l'on ait:

si $f_{nk} \geqslant 0$,

$$\mathsf{E}\{f_{nk}(\zeta)\chi_n|\mathfrak{F}\} \geqslant 0, \quad \mathsf{E}\{rf_{nk}(\zeta)\chi_n|\mathfrak{F}\} = r\mathsf{E}\{f_{nk}(\zeta)\chi_n|\mathfrak{F}\};$$

si $|f_{nk} - f_{nj}| < r$,

$$|\mathsf{E}\{(f_{nk}-f_{nj})\chi_n|\mathfrak{F}\}| \leqslant r,$$

$$\mathsf{E}\{(f_{nk}(\zeta)+f_{nj}(\zeta))\chi_n|\mathfrak{F}\} = \mathsf{E}\{f_{nk}(\zeta)\chi_n|\mathfrak{F}\} + \mathsf{E}\{f_{nj}(\zeta)\chi_n|\mathfrak{F}$$

pour tous les n, k, j et les r rationnels. Par ailleurs

$$\left|\int_F (\mathsf{E}\{f(\zeta)\chi_n|\mathfrak{F}\} - \mathsf{E}\{f_{nk}(\zeta)\chi_n|\mathfrak{F}\})\chi_n(\zeta)\,d\mathsf{P}\right| \leqslant$$

$$\leqslant \int_{F\cap K_n} |f(x) - f_{nk}(x)|\,q(dx),$$

de sorte que si $\|\chi_n(f-f_{nj})\| \to 0$, on a

$$\mathsf{E}\{f(\zeta)\chi_n|\mathfrak{F}\} = \lim \mathsf{E}\{f_{nk_j}(\zeta)\chi_n|\mathfrak{F}\} \pmod{\mathsf{P}}, \tag{13}$$

la limite de droite ne dépendant pas de la suite approximante.

L'espérance mathématique conditionnelle n'étant pas définie sur des ensembles de probabilité nulle, on peut se servir de la relation (13) pour définir $\mathsf{E}\{f(\zeta)\chi_n \mid \mathfrak{F}\}$ dans le cas où f est arbitraire et continue sur K_n, où f_{nkj} est une suite d'éléments d'un réseau partout dense telle que $\|(f - f_{nk_j})\chi_n\| \to 0$. Avec une telle définition des espérances mathématiques conditionnelles, pour tous les $\omega \bar{\in} D_0$, toutes les fonctions $f(x)$ et $g(x)$ continues sur K_n et c et d quelconques réels on a

$$\mathsf{E}\{f(\zeta)\chi_n|\mathfrak{F}\} \geqslant 0 \qquad \text{si } f \geqslant 0,$$

$$\mathsf{E}\{cf(\zeta)\chi_n + dg(\zeta)\chi_n|\mathfrak{F}\} = c\mathsf{E}\{f(\zeta)\chi_n|\mathfrak{F}\} + d\mathsf{E}\{g(\zeta)\chi_n|\mathfrak{F}\},$$

c'est-à-dire $\mathsf{E}\{f(\zeta)\chi_n \mid \mathfrak{F}\}$ est presque sûrement une fonctionnelle linéaire positive sur $\mathscr{C}(K_n)$. D'après un théorème de la théorie de la mesure cette fonctionnelle admet la représentation

$$\mathsf{E}\{f(\zeta)\chi_n|\mathfrak{F}\} = \int_{K_n} f(x)\,q_n(dx, \omega), \quad \omega \bar{\in} D_0,$$

où $q_n(B, \omega)$ sont des mesures définies univoquement sur tous les sous-ensembles boréliens K_n. Pour $B \in \mathfrak{B}$ arbitraire posons

$$q(B, \omega) = \lim q_n(B \cap K_n, \omega), \qquad \omega \bar{\in} D_0,$$
$$q(B, \omega) = q(B), \qquad \omega \in D_0.$$

On vérifie sans peine que $q(B, \omega)$ est une mesure pour ω fixe. En effet, d'abord $q(B, \omega)$ est finiment additive (ceci découle de l'addi-

tivité de q_n) et ensuite, si $B = \bigcup_{n=1}^{\infty} B_n$, $B_{n_1} \cap B_{n_2} = \varnothing$ pour $n_1 \neq \neq n_2$, alors ($\omega \in D_0$)

$$\sum_1^N q(B_k, \omega) = q(\bigcup_1^N B_k, \omega) \leqslant q(\bigcup_1^\infty B_k, \omega) = \lim_{n\to\infty} q_n(\bigcup_1^\infty B_k, \omega) =$$

$$= \lim_{n\to\infty} \sum_1^\infty q_n(B_k \cap K_n, \omega) \leqslant \sum_1^\infty q(B_k, \omega).$$

Lorsque $N \to \infty$, on obtient

$$q(\bigcup_1^\infty B_k, \omega) = \sum_1^\infty q(B_k, \omega),$$

c'est-à-dire $q(B, \omega)$ est dénombrablement additive. L'égalité

$$\int_{Kn} f(x)\, q_n(dx, \omega) = \int_{Kn} f(x)\, q(dx, \omega), \quad \omega \in D_0,$$

implique ($\omega \in D_0$)

$$\int_X f(x)\, q(dx, \omega) = \int_{\bigcup_1^\infty Kn} f(x)\, q(dx, \omega) = \lim \int_{Kn} f(x)\, q_n(dx, \omega)$$

pour toute fonction f non négative $\mathfrak{B}$-mesurable. Donc si f est continue et bornée, alors

$$\mathsf{E}\{f \mid \mathfrak{F}\} = \lim_{n\to\infty} \mathsf{E}\{f\chi_n \mid \mathfrak{F}\} = \int_X f(x)\, q(dx, \omega) \quad (\mathrm{mod}\, \mathsf{P}). \qquad (14)$$

La fonction $f(x)$ $\mathfrak{B}$-mesurable bornée étant susceptible d'être approchée par une suite de fonctions continues $f_n(x)$ de telle sorte que

$$\int_x |f(x) - f_n(x)|\, q(dx) \to 0,$$

de l'inégalité

$$\left| \int_F \mathsf{E}\{f(\zeta) \mid \mathfrak{F}\}\, d\mathsf{P} - \int_F \mathsf{E}\{f_n(\zeta) \mid \mathfrak{F}\}\, d\mathsf{P} \right| \leqslant \int_F |f(x) - f_n(x)|\, q(dx)$$

il suit que (14) a lieu pour toute fonction $\mathfrak{B}$-mesurable bornée. En particulier,

$$\mathsf{P}\{B \mid \mathfrak{F}\} = q(B, \omega) \quad (\mathrm{mod}\, \mathsf{P}). \blacksquare$$

Soient les éléments aléatoires ζ_1 et ζ_2 de $\{Y_1, \mathfrak{B}_1\}$ et $\{Y_2, \mathfrak{B}_2\}$ respectivement, où $\{Y_i, \mathfrak{B}_i\}$ réunissent les conditions du théorè-

me 5. Posons

$$Y^{(1,2)} = Y_1 \times Y_2, \quad \mathfrak{B}^{(1,2)} = \sigma\{\mathfrak{B}_k, \ k = 1, 2\}.$$

La suite $\zeta^{(1,2)} = (\zeta_1, \zeta_2)$ peut être traitée comme un élément aléatoire de $\{Y^{(1,2)}, \mathfrak{B}^{(1,2)}\}$ et $Y^{(1,2)}$ comme un espace polonais. Soient q_i répartition de ζ_i $(i = 1, 2)$, $q^{(1,2)}$ répartition de $\zeta^{(1,2)}$, $q^{(1|2)}$ répartition conditionnelle régulière de ζ_2 par rapport à la tribu $\mathfrak{F}_{\zeta_1}$ engendrée par l'élément aléatoire ζ_1. La répartition $q^{(2|1)}$ étant une fonction $\mathfrak{F}_{\zeta_1}$-mesurable, on a $q^{(2|1)}(B_2, \omega) = q(B_2, \zeta_1)$, où $B_2 \in \mathfrak{B}_2$, et la fonction $q(B_2, y)$ est $\mathfrak{B}_1$-mesurable. La définition des probabilités conditionnelles entraîne

$$\int_{q_1^{-1}(B_1)} q(B_2, \zeta_1)\, d\mathsf{P} = q^{(1,2)}(B_1 \times B_2),$$

où B_1 est un ensemble arbitraire de $\mathfrak{B}_1$ et $\zeta_1 = q_1(\omega)$. Un changement de variables donne

$$q^{(1,2)}(B_1 \times B_2) = \int_{B_1} q(B_2, y_1)\, dy_1$$

ou

$$q^{(1,2)}(B_1 \times B_2) = \int_{Y_1} \chi^{(1)}(B_1, y_1) \int_{Y_2} \chi^{(2)}(B_2, y_2)\, q(dy_2, y_1)\, q_1(dy_1), \tag{15}$$

où $\chi^{(i)}$ sont les indicateurs des ensembles dans l'espace Y_i. La dernière formule entraîne le

Théorème 6.

$$\int_{Y^{(1,2)}} f(y_1, y_2)\, dq^{(1,2)} = \int_{Y_1} \left(\int_{Y_2} f(y_1, y_2)\, q(dy_2, y_1) \right) q_1(dy_1) \tag{16}$$

pour toutes les fonctions non négatives $\mathfrak{B}^{(1,2)}$-mesurables.

En effet, la classe de fonctions f justiciables de la formule (16) est linéaire et fermée pour le passage à la limite sur les suites monotones. Comme en vertu de (15) cette classe renferme les fonctions de la forme $\chi^{(1)} \chi^{(2)}$, elle contiendra leurs combinaisons linéaires. Par ailleurs toute fonction $\mathfrak{B}^{(1,2)}$-mesurable peut être approchée par des suites monotones croissantes de combinaisons linéaires de fonctions de la forme $\chi^{(1)} \chi^{(2)}$. ■

A noter que la formule (16) est vraie pour les fonctions f à signe alterné si seulement l'un des deux membres de (16) a un sens. La formule (16) entraîne le

Corollaire.

$$\mathsf{E}\{f(\zeta_1, \zeta_2) | \mathfrak{F}_{\zeta_1}\} = \int_{Y_2} f(\zeta_1, y_2)\, q(dy_2, \zeta_1). \tag{17}$$

Ce résultat peut être mis sous la forme générale suivante. Soient ζ_k des éléments aléatoires de $\{Y_k, \mathfrak{B}_k\}$, Y_k étant des espaces polonais. Posons

$$Y^{(1, s)} = \prod_{k=1}^{s} Y_k, \quad \mathfrak{B}^{(1, s)} = \sigma\{\mathfrak{B}_k, \quad k = 1, \ldots, s\},$$

$\eta_s = (\zeta_1, \zeta_2, \ldots, \zeta_s)$, q_k est la répartition des éléments ζ_k dans $\{Y_k, \mathfrak{B}_k\}$, $q^{(s)} = q^{(s)}(B_s, \zeta_1, \ldots, \zeta_{s-1})$, la répartition conditionnelle régulière de l'élément ζ_s par rapport à une tribu $\mathfrak{F}_{\eta_{s-1}} = \mathfrak{F}_{(\zeta_1, \zeta_2, \ldots, \zeta_{s-1})}$. En appliquant (17) à la formule

$$\mathsf{E}\{f \mid \mathfrak{F}_{\zeta_1}\} = \mathsf{E}\{\ldots\{\mathsf{E}\{f \mid \mathfrak{F}_{\eta_{n-1}}\} \mid \mathfrak{F}_{\eta_{n-2}}\}\ldots \mid \mathfrak{F}_{\eta_1}\},$$

on obtient

$$\mathsf{E}\{f(\zeta_1, \ldots, \zeta_n) \mid \mathfrak{F}_{\zeta_1}\} =$$

$$= \int_{Y_2} \ldots \int_{Y_{n-1}} \left(\int_{Y_n} f(\zeta_1, y_2, \ldots, y_n)\, q^{(n)}(dy_n, \zeta_1, y_2, \ldots, y_{n-1}) \right) \times$$
$$\times\, q^{(n-1)}(dy_{n-1}, \zeta_1, y_2, \ldots, y_{n-2}) \ldots q^{(2)}(dy_2, \zeta_1), \quad (18)$$

$$\mathsf{E}\{f(\zeta_1, \ldots, \zeta_n\} =$$

$$= \int_{Y_1} \ldots \int_{Y_{n-1}} \left(\int_{Y_n} f(y_1, \ldots, y_n)\, q^{(n)}(dy_n, y_1, \ldots, y_{n-1}) \right) \times$$
$$\times\, q^{(n-1)}(dy_{n-1}, y_1, \ldots, y_{n-2}) \ldots q^{(2)}(dy_2, y_1)\, q_1(dy_1). \quad (19)$$

§ 4. Indépendance

Soit $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ un espace probabilisé fixe. Par événement on entendra, sauf mention expresse du contraire, des parties $\mathfrak{S}$-mesurables de Ω.

Deux événements A et B sont *indépendants* si

$$\mathsf{P}(A \cap B) = \mathsf{P}(A)\, \mathsf{P}(B).$$

S $\mathsf{P}(B) > 0$, cette condition équivaut à $\mathsf{P}(A \mid B) = \mathsf{P}(A)$. La définition implique aussitôt:

a) Ω et A, où A est un événement quelconque, sont indépendants;

b) si $\mathsf{P}(N) = 0$, A quelconque, N et A sont indépendants;

c) si A et B_i ($i = 1, 2$) sont indépendants et $B_1 \supset B_2$, A et $B_1 \setminus B_2$ sont indépendants. En particulier, A et $\overline{B_1}$ sont indépendants;

d) si A et B_i sont indépendants, $i = 1, 2, \ldots, n$, et $B_1, B_2, \ldots$ $\ldots, B_n$ sont deux à deux incompatibles, alors A et $\bigcup_1^n B_i$ sont également indépendants.

A signaler que si aucune restriction n'est faite sur l'incompatibilité deux à deux des événements B_i, la dernière proposition n'a généralement pas lieu;

e) A ne dépend pas de A si et seulement si $\mathsf{P}(A) = 0$ où $\mathsf{P}(A) = 1$.

Soient I un ensemble, $\{\mathfrak{M}_i, i \in I\}$ un ensemble de classes d'événements.

DÉFINITION. *Des classes d'événements de* $\{\mathfrak{M}_i, i \in I\}$ *sont indépendantes (ou indépendantes dans leur ensemble) si pour* $i_1, \ldots, i_n$ $(i_k \in I)$ *deux à deux inégaux quelconques et* A_{i_k} *quelconques,* $A_{i_k} \in \mathfrak{M}_{i_k}, k = 1, 2, \ldots, n$, *on a*

$$\mathsf{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_n}) = \mathsf{P}(A_{i_1}) \mathsf{P}(A_{i_2}) \ldots \mathsf{P}(A_{i_n}).$$

On remarquera que s'agissant d'un ensemble infini de classes d'événements, la définition de l'indépendance revient à exiger qu'un sous-ensemble fini quelconque de classes d'événements soit composé de classes indépendantes d'événements.

Une classe d'événements est par définition une π-classe si elle est fermée pour l'intersection des événements, c'est-à-dire $A \in \mathfrak{M}$ et $B \in \mathfrak{M}$ implique $A \cap B \in \mathfrak{M}$.

THÉORÈME 1. *Soit* $\{\mathfrak{M}_i, i \in I\}$ *une collection de* π*-classes d'événements indépendantes. Les plus petites tribus* $\sigma\{\mathfrak{M}_i\}$ *sont indépendantes.*

Démonstration. On peut se limiter à un nombre fini de classes $\mathfrak{M}_1, \ldots, \mathfrak{M}_n$. Il suffit de prouver que si une classe, par exemple $\mathfrak{M}_1$, est remplacée par $\sigma\{\mathfrak{M}_1\}$, la nouvelle suite de classes d'événements sera également indépendante.

Soit $\mathfrak{A}$ la classe de tous les événements ne dépendant pas de $\mathfrak{M}_2, \ldots, \mathfrak{M}_n$. Par définition $\mathfrak{M}_1 \subset \mathfrak{A}$ et $\mathfrak{A}$ possède les propriétés suivantes: elle est fermée pour l'union de toute suite dénombrable d'événements disjoints et pour la différence $B_2 \setminus B_1$ pourvu que $B_2 \supset B_1$. Le théorème 1 découle maintenant de la proposition suivante.

THÉORÈME 2. *Si une classe d'événements* $\mathfrak{A}$ *contient une* π*-classe* $\mathfrak{M}$ *et possède les propriétés suivantes:*

a) $A_1 \subset A_2$, $A_i \in \mathfrak{A}$, $i = 1, 2$, *implique* $A_2 \setminus A_1 \in \mathfrak{A}$;

b) $A_n \in \mathfrak{A}, n = 1, 2, \ldots, A_n \cap A_m = \varnothing$ *pour* $n \neq m$ *implique* $\bigcup_1^\infty A_n \in \mathfrak{A}$,

alors $\mathfrak{A} \supset \sigma \{\mathfrak{M}\}$.

D é m o n s t r a t i o n. Désignons par $\mathfrak{A}_1$ la plus petite classe d'événements renfermant $\mathfrak{M}$ et possédant les propriétés a) et b) ($\mathfrak{A}_1$ est l'intersection de toutes les classes contenant $\mathfrak{M}$ et possédant les propriétés a) et b)). Prouvons que $\mathfrak{A}_1 = \sigma \{\mathfrak{M}\}$. Désignons par $\mathfrak{A}_1 (B)$ la classe de tous les événements $A \in \mathfrak{A}_1$ tels que $A \cap \cap B \in \mathfrak{A}_1$. On vérifie sans peine que $\mathfrak{A}_1 (B)$ possède les propriétés a) et b). D'autre part, si $B \in \mathfrak{M}$, alors $\mathfrak{A}_1 (B) \supset \mathfrak{M}$, puisque $\mathfrak{M}$ est une π-classe. Donc $\mathfrak{A}_1 (B) = \mathfrak{A}_1$ (si $B \in \mathfrak{M}$). Or cela signifie que $\mathfrak{A}_1 (A) \supset \mathfrak{M}$ quel que soit $A \in \mathfrak{A}_1$, c'est-à-dire $\mathfrak{A}_1 (A) = \mathfrak{A}_1$ quel que soit $A \in \mathfrak{A}_1$. Donc $\mathfrak{A}_1$ est une π-classe. Or une π-classe d'événements possédant à la fois les propriétés a) et b) est une tribu. En définitive $\mathfrak{A}_1 = \sigma \{\mathfrak{M}\}$ et $\mathfrak{A} \supset \sigma \{\mathfrak{M}\}$. ■

THÉORÈME 3. *Soit* $\{\mathfrak{M}_i, i \in I\}$ *un ensemble de* π-*classes d'événements indépendantes*, $I = I_1 \cup I_2$ $(I_1 \cap I_2 = \varnothing)$, $\mathfrak{B}_k = \sigma \{\mathfrak{M}_i, \; i \in I_k\}$, $k = 1, 2$. *Alors* $\mathfrak{B}_1$ *et* $\mathfrak{B}_2$ *sont indépendantes.*

En vertu du théorème 1 on retiendra seulement l'hypothèse que $\mathfrak{M}_i$ est une tribu. Considérons les classes $\mathfrak{A}_k$ $(k = 1, 2)$ composées de tous les événements de la forme $A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_n}$, n est quelconque, $i_r \in I_k$. Ces classes sont fermées pour les intersections, $\mathfrak{A}_k$ contient tous les $\mathfrak{M}_i$, $i \in I_k$ et $\mathfrak{A}_1$ et $\mathfrak{A}_2$ sont indépendantes. D'après le théorème 2 $\sigma \{\mathfrak{A}_1\} = \sigma \{\mathfrak{M}_i, \; i \in I_1\}$ et $\sigma \{\mathfrak{A}_2\} = \sigma \{\mathfrak{M}_i, \; i \in I_2\}$ sont indépendantes.

COROLLAIRE. *Si l'on partage* I *en un nombre quelconque de parties deux à deux disjointes*, $I = \bigcup_{k \in Q} I_k$, *les tribus* $\mathfrak{B}_k = \sigma \{\mathfrak{M}_i, \; i \in I_k\}$, $k \in Q$, *sont indépendantes dans leur ensemble.*

Eléments aléatoires indépendants. Soit $\zeta_i = f_i (\omega)$ élément aléatoire de $\{X_i, \mathfrak{B}_i\}$, $i \in I$.

DÉFINITION. *On dit que des éléments aléatoires* $\{\zeta_i, \; i \in I\}$ *sont indépendants* (*indépendants dans leur ensemble*) *si pour tout* n *et* $B_k \in \in \mathfrak{B}_{i_k}$, $i_k \in I$, *on a*

$$\mathsf{P}\left(\bigcap_{k=1}^{n} \{\zeta_{i_k} \in B_k\}\right) = \prod_{k=1}^{n} \mathsf{P}\{\zeta_{i_k} \in B_k\}.$$

La définition de l'indépendance de familles d'ensembles d'éléments aléatoires est plus générale. Considérons une famille d'ensembles $\{\zeta_i^\mu, \; i \in I_\mu\}$, $\mu \in M$, d'éléments aléatoires à valeurs dans $\{X_i^\mu, \mathfrak{B}_i^\mu\}$.

Définition. *Les ensembles d'éléments aléatoires* $\{\zeta_i^\mu,\ i \in I_\mu\}$ $(\mu \in M)$ *sont indépendants (mutuellement indépendants) si le sont les classes d'événements* $\{\mathfrak{M}_\mu,\ \mu \in M\}$, *où* $\mathfrak{M}_\mu$ *est composé de tous les événements de la forme*

$$\bigcap_{k=1}^{n} \{\zeta_{i_k}^\mu \in B_{i_k}^\mu\}, \quad n=1,\ 2,\ \ldots,\ i_k \in I_\mu,\ B_{i_k}^\mu \in \mathfrak{B}_{i_k}^\mu.$$

Supposons que $\sigma\,\{\zeta_i^\mu,\ i \in I_\mu\} = \mathfrak{F}_\mu$ est la tribu engendrée par l'ensemble des éléments aléatoires ζ_i^μ, $i \in I_\mu$, c'est-à-dire la plus petite tribu par rapport à laquelle sont mesurables tous les éléments aléatoires ζ_i^μ, $i \in I_\mu$, μ étant fixe.

Théorème 4. *Les ensembles d'éléments aléatoires* $\{\zeta_i^\mu,\ i \in I_\mu\}$ $(\mu \in \in M)$ *sont indépendants si et seulement si le sont les tribus* $\mathfrak{F}_\mu$, $\mu \in M$.

La démonstration découle du fait que les classes d'événements intervenant dans la définition de l'indépendance d'ensembles d'éléments aléatoires sont des π-classes et du théorème 1.

Corollaire. *Soient* $(\zeta_1^\mu, \ldots, \zeta_s^\mu)$, $\mu \in M$, *un ensemble de suites indépendantes d'éléments aléatoires,* $g_\mu(x_1, \ldots, x_{s_\mu})$ *des fonctions* $\sigma\,\{\mathfrak{B}_k^\mu,\ k = 1, \ldots, s_\mu\}$*-mesurables* $(\mu \in M)$. *Les éléments aléatoires*

$$\xi_\mu = g_\mu(\zeta_1^\mu, \ldots, \zeta_{s_\mu}^\mu), \quad \mu \in M,$$

sont alors mutuellement indépendants.

Remarque. Les éléments aléatoires $\{\xi_\mu,\ \mu \in M\}$ sont mutuellement indépendants si et seulement si quels que soient n, μ_1, μ_2, ..., μ_n de M et $a_1, \ldots, a_n$ réels, on a

$$\mathsf{P}\{\xi_{\mu_1} < a_1, \ldots, \xi_{\mu_n} < a_n\} = \prod_{k=1}^{n} \mathsf{P}\{\xi_{\mu_k} < a_k\}.$$

La condition nécessaire est triviale. La condition suffisante découle du fait que la tribu engendrée par les événements de la forme $\{\xi_\mu < a\}$ est confondue avec la tribu $\{\xi_\mu \in B,\ B \in \mathfrak{B}^1\}$, $\mathfrak{B}^1$ est la tribu des boréliens de la droite.

Soient ζ_k, $k=1, 2, \ldots, n$, suite d'éléments aléatoires indépendants (sur $\{X_k, \mathfrak{B}_k\}$ respectivement), q_k répartition de ζ_k sur $\mathfrak{B}_k$, $q^{(1,n)}$ répartition conjointe de la suite $(\zeta_1, \ldots, \zeta_n)$ dans $\{\prod_{k=1}^{n} X_k,\ \sigma\{\mathfrak{B}_k,\ k=1, \ldots, n\}\}$. Par définition de l'indépendance on a

$$q^{(1,n)}(B_1 \times B_2 \times \ldots \times B_n) = \prod_{k=1}^{n} q_k(B_k), \quad B_k \in \mathfrak{B}_k. \tag{1}$$

La réciproque est évidente, c'est-à-dire si est réalisé (1) pour tous les $B_k \in \mathfrak{B}_k$: les éléments $\{\zeta_k, k = 1, \ldots, n\}$ sont indépendants.

Soient $g(x_1, x_2)$ fonction $\sigma\{\mathfrak{B}_i, i = 1, 2\}$-mesurable, ζ_1, ζ_2 éléments aléatoires indépendants et $\mathsf{E}g(\zeta_1, \zeta_2) < \infty$. Un changement de variables et le théorème de Fubini nous apprennent que $\mathsf{E}g(\zeta_1, \zeta_2)$ est une fonction $\mathfrak{B}_1$-mesurable, finie pour q_1-presque tous les x_1 et

$$\mathsf{E}g(\zeta_1, \zeta_2) = \int_{X_1} q_1(dx_1) \int_{X_2} g(x_1, x_2)\, q_2(dx_2). \tag{2}$$

De là découle la formule

$$\mathsf{E}g(\zeta_1)\, g(\zeta_2) = \mathsf{E}g(\zeta_1)\, \mathsf{E}g(\zeta_2), \tag{3}$$

qui est valable si $\mathsf{E}g(\zeta_k)$ sont finies.

La proposition suivante renforce la précédente.

THÉORÈME 5. *Si une variable aléatoire ξ à espérance mathématique finie est indépendante par rapport à une tribu $\mathfrak{F}$, on a*

$$\mathsf{E}\{\xi \mid \mathfrak{F}\} = \mathsf{E}\xi.$$

Démonstration. L'indépendance de la variable aléatoire ξ par rapport à la tribu $\mathfrak{F}$ signifie que les tribus $\mathfrak{F}$ et $\mathfrak{F}_\xi = \sigma\{\xi\}$ sont indépendantes. Donc pour tout $F \in \mathfrak{F}$ les variables aléatoires ξ et $\chi(F)$ sont indépendantes. Par conséquent

$$\int_F \xi\, d\mathsf{P} = \mathsf{E}\chi(F)\, \mathsf{E}\xi = \int_F (\mathsf{E}\xi)\, d\mathsf{P}.$$

Etant constante, $E\xi$ est $\mathfrak{F}$-mesurable. Donc

$$\mathsf{E}\xi = \mathsf{E}\{\xi \mid \mathfrak{F}\}. \blacksquare$$

Loi de tout ou rien. Soit A_n, $n = 1, 2, \ldots$, une suite d'événements.

THÉORÈME 6. (Théorème de Borel-Cantelli). *Si* $\sum_{n=1}^{\infty} \mathsf{P}(A_k) < \infty$, *l'événement* $\overline{\lim}\, A_n = \{A_n$ *infiniment fréquent*$\}$ *possède la probabilité* 0. *Si les événements* A_n, $n = 1, 2, \ldots$, *sont indépendants, la probabilité de l'événement* $\overline{\lim}\, A_n$ *est égale à* 0 *ou* 1 *selon que la série* $\sum_{n=1}^{\infty} \mathsf{P}(A_n)$ *est convergente ou divergente.*

Démonstration.

a) Comme $\overline{\lim}\, A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$, on a

$$\mathsf{P}\{A_n \text{ inf.fr.}\} = \lim_{n\to\infty} \mathsf{P}\Big(\bigcup_{k=n}^{\infty} A_k\Big) \leqslant \lim_{n\to\infty} \sum_{k=n}^{\infty} \mathsf{P}(A_k) = 0,$$

ce qui démontre la première partie du théorème.

b) Supposons maintenant que les événements A_n sont indépendants. Il faut seulement prouver que si

$$\sum_{n=1}^{\infty} \mathsf{P}(A_n) = \infty,$$

alors

$$\mathsf{P}\{\overline{\lim} A_n\} = 1.$$

Soit

$$A^* = \overline{\lim} A_n,$$

il vient

$$\Omega \setminus A^* = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} (\Omega \setminus A_k)$$

et

$$\mathsf{P}(\Omega \setminus A^*) = \lim_{n\to\infty} \mathsf{P}\Big(\bigcap_{k=n}^{\infty} (\Omega \setminus A_k)\Big) = \lim_{n\to\infty} \prod_{k=n}^{\infty} \mathsf{P}(\Omega \setminus A_k) = $$

$$= \lim_{n\to\infty} \prod_{k=n}^{\infty} (1 - \mathsf{P}(A_k)) = 0$$

vu la divergence de la série $\sum_{k=1}^{\infty} \mathsf{P}(A_k)$. ∎

Soit maintenant une suite quelconque de tribus $\mathfrak{F}_n$, $n = 1, 2, \ldots$, indépendantes. En vertu du théorème de Borel-Cantelli l'événement $A^* = \overline{\lim} A_n$, où A_n est une suite quelconque telle que $A_n \in \mathfrak{F}_n$, possède la probabilité 0 ou 1. Ce résultat peut être généralisé à des événements quelconques engendrés par l'ensemble de toutes les tribus $\mathfrak{F}_n$, $n = 1, 2, \ldots$, et ne dépendant pas d'une suite finie quelconque de tribus $\mathfrak{F}_1, \mathfrak{F}_2, \ldots, \mathfrak{F}_n$. Précisons cette proposition.

Soit $\mathfrak{B}_k = \sigma\{\mathfrak{F}_j, j = k, k+1, \ldots\}$, $\mathfrak{B}_k$ forment une suite monotone décroissante de tribus. Leur intersection $\mathfrak{B} = \bigcap_{k=1}^{\infty} \mathfrak{B}_k$ est aussi une tribu. Posons par définition

$$\mathfrak{B} = \overline{\lim}\, \mathfrak{F}_n = \bigcap_{k=1}^{\infty} \sigma\{\mathfrak{F}_j, j = k, k+1, \ldots\}.$$

De toute évidence, la tribu $\overline{\lim}\, \mathfrak{F}_n$ ne change pas si l'on remplace un nombre fini quelconque de tribus $\mathfrak{F}_1, \mathfrak{F}_2, \ldots, \mathfrak{F}_n$ par d'autres.

THÉORÈME 7 (loi générale de tout ou rien de Kolmogorov). *Si* $\mathfrak{F}_n$, $n = 1, 2, \ldots$, *sont des tribus mutuelle-*

ment indépendantes, tout événement appartenant à $\overline{\lim}\mathfrak{F}_n$ *possède la probabilité* 0 *ou* 1.

En effet, soit $A \in \overline{\lim}\mathfrak{F}_n$. Alors $A \in \mathfrak{B}_k$ quel que soit k. Donc A et $\sigma\{\mathfrak{F}_1, \ldots, \mathfrak{F}_{k-1}\}$ sont indépendants. Par conséquent A et $\sigma\{\mathfrak{F}_1, \ldots, \mathfrak{F}_n, \ldots\}$ le sont. Comme $A \in \sigma\{\mathfrak{F}_k, k = 1, 2, \ldots\}$, il suit que A ne dépend pas de A. Or ceci n'est possible que si $\mathsf{P}(A) = 0$ ou $\mathsf{P}(A) = 1$. ■

Théorème 8. *Soient* $\{\zeta_n, n = 1, 2, \ldots\}$ *une suite d'éléments aléatoires indépendants dans un espace métrique fixé* $\{X, \mathfrak{B}\}$, $\mathfrak{F}_n$ *la tribu engendrée par* ζ_n *et* $\mathfrak{B}_n = \sigma\{\mathfrak{F}_k, k = n, n+1, \ldots\}$. *Alors*

a) *la limite de la suite* $\{\zeta_n, n = 1, 2, \ldots\}$ *existe avec la probabilité* 1 *ou* 0;

b) *si* X *est polonais, la limite de la suite* $\{\zeta_n, n = 1, 2, \ldots\}$, *si elle existe, est presque sûrement constante*;

c) *si* $z = f(x_1, x_2, \ldots, x_n, \ldots)$ *est une fonction d'un nombre infini de variables* $x_n \in X$, $n = 1, 2, \ldots$, *et* $f(\xi_1, \ldots, \xi_n, \ldots)$ *est* $\mathfrak{B}_n$*-mesurable quel que soit* n, *alors elle est presque sûrement constante.*

Démonstration.

a) Si $\rho(x, y)$ est la distance dans X, alors l'ensemble des points de convergence de ζ_n s'écrit

$$D = \bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{n', n'' \geq n} \left\{\omega : |\zeta_{n'} - \zeta_{n''}| < \frac{1}{k}\right\}.$$

Comme les événements

$$A_n = \bigcap_{n', n'' \geq n} \left\{|\zeta_{n'} - \zeta_{n''}| < \frac{1}{k}\right\}$$

sont monotones croissants, on a

$$\bigcup_{n=1}^{\infty} A_n \in \mathfrak{B}_m$$

quel que soit m, de sorte que $D \in \mathfrak{B}_m$ pour tout m et l'on peut appliquer la loi de tout ou rien générale.

b) Soit F un fermé, $F \subset X$; appelons F_k l'ouvert $F_k = \{x : \rho(x, F) < \frac{1}{k}\}$. L'événement $D \cap \{\lim \zeta_n \in F\}$ s'écrit alors

$$D \cap \{\lim \zeta_n \in F\} = D \cap \left[\bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k \geq n} \{\zeta_n \in F_k\}\right],$$

il appartient à $\mathfrak{B}_m$, $m = 1, 2, \ldots$, en vertu des mêmes raisonnements qui ont présidé à la démonstration de a). Donc, pour tout ensemble fermé F on a $\mathsf{P}\{\lim \zeta_n \in F\} = 0$ ou 1. Or la classe d'ensembles $\mathfrak{B}$ pour laquelle cette proposition a lieu est une tribu. Donc $\mathsf{P}\{\lim \zeta_n \in B\} = 0$ ou 1 pour tout $B \in \mathfrak{B}$. Si l'espace X est polonais, on déduit sans peine que la mesure q induite sur $\mathfrak{B}$ par l'élément

aléatoire $\lim \zeta_n$ est concentrée sur un seul atome. En effet, étant donné que $q(X) = 1$, il existe une boule S_1 de rayon 1 telle que $q(S_1) = 1$ (si une telle boule n'existait pas, toutes les boules de X de rayon 1 auraient une mesure nulle, ce qui est impossible puisque X est recouvert d'un nombre dénombrable de telles boules). De façon analogue il existe une boule S_2 de rayon $\frac{1}{2}$, $S_2 \subset S_1$ et $q(S_2) = 1$. En poursuivant cette procédure on obtient une suite de boules emboîtées S_n de rayons $\frac{1}{n}$ et de mesure 1. Ces boules ne possèdent que le point x en commun et $q\{x\} = \lim q(S_n) = 1$.

c) Par hypothèse les événements $A = \{\omega : f(\zeta_1, \zeta_2, \ldots, \zeta_n) < a\} \in \mathfrak{B}_n$. Par suite $A \in \overline{\lim} \mathfrak{F}_n$ et A possèdent la probabilité 0 ou 1. Donc la fonction de répartition de la variable aléatoire $\zeta = f(\zeta_1, \ldots, \zeta_n, \ldots)$ ne prend que la valeur 0 ou 1 et ζ est constante presque sûrement. ■

CHAPITRE III

SUITES ALÉATOIRES

§ 1. Martingales

Définition et propriétés élémentaires. Une martingale est une famille de variables aléatoires $\xi(t)$, $t \in T$ (T est l'ensemble des réels), sans mémoire, c'est-à-dire que les espérances mathématiques conditionnelles des accroissements $\xi(t_2) - \xi(t_1)$ $(t_1 < t_2)$ pour $\xi(s)$ donnés, $s \leqslant t$, sont nulles indépendamment de ces valeurs. Si les espérances mathématiques conditionnelles sont non négatives (resp. non positives), $\xi(t)$ est appelée submartingale (resp. supermartingale).

Précisons cette définition. Dans ce paragraphe on étudiera essentiellement des suites de variables aléatoires. Nous allons tout d'abord donner la définition générale. Soient $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ un espace probabilisé, T l'ensemble des réels. Nous traiterons $t \in T$ comme les instants de réalisation des expériences. L'ensemble des observables avant l'instant t sera désigné par $\mathfrak{F}_t$. Naturellement $\mathfrak{F}_{t_1} \subset \mathfrak{F}_{t_2}$ si $t_1 < t_2$. Considérons une famille de variables aléatoires $\xi(t)$, $t \in T$, telles que les valeurs de $\xi(t)$ soient exactement définies par un ensemble d'expériences effectuées à des instants s, $s \leqslant t$. En d'autres termes les variables $\xi(t)$ doivent être $\mathfrak{F}_t$-mesurables pour tout $t \in T$. Les définitions suivantes nous aideront à décrire formellement cette situation.

DÉFINITION. *On appellera flot de tribus une famille monotone croissante de tribus* $\{\mathfrak{F}_t, t \in T\}$ ($\mathfrak{F}_{t_1} \subset \mathfrak{F}_{t_2}$ *si* $t_1 < t_2$, $\mathfrak{F}_t \subset \mathfrak{S}$). *Une famille de variables aléatoires* $\{\xi(t), t \in T\}$ *est adaptée au flot* $\{\mathfrak{F}_t, t \in T\}$ *si* $\xi(t)$ *est* $\mathfrak{F}_t$*-mesurable pour tout* $t \in T$.

DÉFINITION. *Une famille* $\{\xi(t), \mathfrak{F}_t, t \in T\}$ *dont les variables aléatoires* $\xi(t)$ *sont adaptées à un flot de tribus* $\{\mathfrak{F}_t, t \in T\}$ *et* $\mathsf{E}\,|\xi(t)| < \infty$, $t \in T$, *est une martingale si pour* $s < t$, $s, t \in T$

$$\mathsf{E}\,\{\xi(t) \mid \mathfrak{F}_s\} = \xi(s), \tag{1}$$

une submartingale si

$$\mathsf{E}\,\{\xi(t) \mid \mathfrak{F}_s\} \geqslant \xi(s), \tag{2}$$

une supermartingale si

$$\mathsf{E}\,\{\xi(t)\mid\mathfrak{F}_s\}\leqslant\xi(s). \tag{3}$$

De la définition il résulte que la tribu $\sigma\{\xi(s),\ s\leqslant t,\ s\in T\}$ appartient à $\mathfrak{F}_t$. Dans de nombreux cas on admet que $\mathfrak{F}_t = \sigma\{\xi(s),\ s\leqslant t,\ s\in T\}$, mais souvent par $\mathfrak{F}_t$ on entendra une tribu plus grande. Dans les cas où l'on sait de quel flot de tribus il est question, on appellera martingale (ou submartingale) la famille $\{\xi(t),\ t\in T\}$.

Les propriétés établies pour les submartingales sont valables au signe près pour les supermartingales.

La définition des espérances mathématiques conditionnelles entraîne que la famille $\{\xi(t),\ \mathfrak{F}_t,\ t\in T\}$ est une martingale (resp. submartingale) si et seulement si elle est attachée au flot $\mathfrak{F}_t$, $t\in T$, $\mathsf{E}\,|\xi(t)|<\infty$ et quels que soient s, t; $A\in\mathfrak{F}_s$ $(s<t,\ s,t\in T)$, on a

$$\int_A \xi(s)\,d\mathsf{P}=\int_A \xi(t)\,d\mathsf{P} \tag{4}$$

$$\left(\text{resp. } \int_A \xi(s)\,d\mathsf{P}\leqslant\int_A \xi(t)\,d\mathsf{P}\right). \tag{5}$$

En particulier, $\mathsf{E}\xi(t)=\text{const}$ pour une martingale et $\mathsf{E}\xi(s)\leqslant\mathsf{E}\xi(t)$ $(s<t)$ pour une submartingale.

Il est évident que si $\xi(t)$ et $\eta(t)$ sont deux submartingales, $a>0$, $b>0$, la combinaison linéaire $a\xi(t)+b\eta(t)$ est également submartingale. D'un autre côté la combinaison linéaire de deux martingales est toujours une martingale.

THÉORÈME 1.

a) *Si* $\{\xi(t),\ \mathfrak{F}_t,\ t\in T\}$ *est une martingale,* $f(x)$ *une fonction convexe continue et* $\mathsf{E}\,|f(\xi(t))|<\infty$, *alors* $\{f(\xi(t)),\ \mathfrak{F}_t,\ t\in T\}$ *est une submartingale.*

b) *Si* $\{\xi(t),\ \mathfrak{F}_t,\ t\in T\}$ *est une submartingale et* $f(x)$ *une fonction convexe continue monotone non décroissante,* $\mathsf{E}\,|f(\xi(t))|<\infty$, *alors* $\{f(\xi(t)),\ \mathfrak{F}_t,\ t\in T\}$ *est également submartingale.*

c) *Si* $\xi(t)$, $\eta(t)$, $t\in[0, T]$ *sont des submartingales,* $\xi(t)\vee\eta(t)$ *est également submartingale.*

Ces assertions découlent aisément de l'inégalité de Jensen. En effet, si f est convexe et $\xi(t)$ martingale, $s<t$, on a

$$\mathsf{E}\{f(\xi(t))\mid\mathfrak{F}_s\}\geqslant f(\mathsf{E}\{\xi(t)\mid\mathfrak{F}_s\})=f(\xi(s)),$$

c'est-à-dire $f(\xi(t))$ est une submartingale. Si $\xi(t)$ est submartingale et la fonction $f(x)$ convexe et monotone non décroissante, alors

$$\mathsf{E}\{f(\xi(t))\mid\mathfrak{F}_s\}\geqslant f(\mathsf{E}\{\xi(t)\mid\mathfrak{F}_s\})\geqslant f(\xi(s)).$$

Soit $\zeta(t) = \xi(t) \vee \eta(t)$. Le processus $\zeta(t)$ est attaché au flot $\mathfrak{F}_t$ et est intégrable. Pour tout $A \in \mathfrak{F}_s$, où $s \in T$, $s < t$, et eu égard au fait que les événements $A_1 = A \cap \{\xi(s) < \eta(s)\}$ et $A_2 = A \cap \cap \{\xi(s) \geqslant \eta(s)\}$ sont $\mathfrak{F}_s$-mesurables, on a

$$\int_A \zeta(s)\, d\mathsf{P} = \int_{A_1} \eta(s)\, d\mathsf{P} + \int_{A_2} \xi(s)\, d\mathsf{P} \leqslant$$

$$\leqslant \int_{A_1} \eta(t)\, d\mathsf{P} + \int_{A_2} \xi(t)\, d\mathsf{P} \leqslant \int_A \zeta(t)\, d\mathsf{P}.$$

Donc $\zeta(t)$ est une submartingale. ■

COROLLAIRE.

1) *Si* $\xi(t)$, $t \in T$, *est une submartingale, alors* $a \vee \xi(t)$, $t \in T$, *l'est également* (a *est une constante*).

2) *Si* $\xi(t)$, $t \in T$, *est une martingale*, $|\xi(t)|$ *est une submartingale et* $\mathsf{E}\,|\xi(t)|$ *est une fonction monotone non décroissante. Si* $p > 1$ *et* $\mathsf{E}\,|\xi(t)|^p < \infty$, *alors* $|\xi(t)|^p$ *est une submartingale. Si* $\xi(t)$ *est une submartingale, alors* $\xi^+(t)$ *l'est aussi. De plus, si* $p > 1$ *et* $\mathsf{E}\,(\xi^+)^p < \infty$, *alors* $(\xi^+)^p$ *est également submartingale.*

EXEMPLES. a) Soient ζ_k, $k = 0, 1, \ldots, n, \ldots$, des variables aléatoires à espérances mathématiques finies, $\mathfrak{F}_n = \sigma\{\zeta_1, \ldots, \zeta_n\}$, $\xi_n = \zeta_1 + \zeta_2 + \ldots + \zeta_n$. Si $\mathsf{E}\zeta_n = 0$ $(\mathsf{E}\zeta_n \geqslant 0)\ \forall n$, alors $\{\xi_n, \mathfrak{F}_n, n = 1, 2, \ldots\}$ est une martingale (submartingale).

En effet, pour $m < n$

$$\mathsf{E}\{\xi_n \mid \mathfrak{F}_m\} = \mathsf{E}\{\xi_n - \xi_m \mid \mathfrak{F}_m\} + \xi_m = \xi_m + \sum_{k=m+1}^{n} \mathsf{E}\xi_k.$$

De façon analogue, si $\mathsf{E}\zeta_n = 1$ $(\mathsf{E}\zeta_n \geqslant 1)\,\forall n$, la suite $\{\xi_n, \mathfrak{F}_n, n = 1, 2, \ldots\}$, où $\xi_n = \prod_{k=1}^{n} \zeta_k$, est martingale (submartingale). En effet $(m < n)$,

$$\mathsf{E}\{\xi_n \mid \mathfrak{F}_m\} = \xi_m \mathsf{E}\left\{\prod_{m+1}^{n} \zeta_k \mid \mathfrak{F}_m\right\} = \xi_m \prod_{m+1}^{n} \mathsf{E}\zeta_k.$$

b) Soient $\{\mathfrak{F}_n, n = 1, 2, \ldots\}$ un flot de tribus, $\mathfrak{F}_n \subset \mathfrak{S}$, $\mathfrak{S}$ une tribu de parties de Ω, P et Q deux mesures probabilistes sur $\mathfrak{S}$, P_n et Q_n les restrictions des mesures P et Q à $\mathfrak{F}_n$. Supposons que Q_n est absolument continue par rapport à P_n $\forall n$. Considérons sur l'espace probabilisé $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ la suite de variables aléatoires ξ_n :

$$\xi_n = \frac{d\mathsf{Q}_n}{d\mathsf{P}_n}, \quad n = 1, 2, \ldots,$$

où $\frac{dQ_n}{dP_n}$ est la dérivée au sens de Radon-Nikodym de la mesure Q_n par rapport à P_n. Pour $A \in \mathfrak{F}_m$, $m < n$, on a alors

$$\int_A \xi_n d\mathsf{P} = Q_n(A) = Q_m(A) = \int_A \xi_m d\mathsf{P},$$

c'est-à-dire $\{\xi_n, \mathfrak{F}_n, \mathsf{P}\}$ est une martingale.

Soit en particulier $\eta_1, \ldots, \eta_n, \ldots$ une suite de vecteurs aléatoires à valeurs dans $\mathscr{R}^d$, $f_n(x_1, \ldots, x_n)$, $x_k \in \mathscr{R}^d$, $k = 1, \ldots, n$, la densité de probabilité conjointe des vecteurs $(\eta_1, \ldots, \eta_n)$ par rapport à une mesure σ-finie q_n définie sur la tribu des boréliens de de $\mathscr{R}^{dn}$. Supposons que $g(x_1, \ldots, x_n)$ est une densité par rapport à q_n telle que

$$\int_A f_n\, dq_n = 0 \Rightarrow \int_A g_n\, dq_n = 0.$$

La suite

$$\frac{g_n(\eta_1, \ldots, \eta_n)}{f_n(\eta_1, \ldots, \eta_n)}, \quad n = 1, 2, \ldots,$$

est une martingale par rapport au flot $\mathfrak{F}_n$, où $\mathfrak{F}_n = \sigma\{\eta_1, \ldots, \eta_n\}$.

Instants markoviens. Soit $\{\mathfrak{F}_t, t \in T\}$ un flot de tribus.

DÉFINITION. *On appelle instant markovien τ sur un flot de tribus $\{\mathfrak{F}_t, t \in T\}$ une variable aléatoire prenant ses valeurs sur T ou la valeur $+\infty$ et telle que $\{\tau \leqslant t\} \in \mathfrak{F}_t$ pour tout $t \in T$.*

Si τ est un instant markovien, alors quel que soit $t \in T$ les événements $\{\tau > T\}$, $\{\tau < T\}$, $\{\tau = T\}$ sont $\mathfrak{F}_t$-mesurables.

Associons à tout instant markovien τ une tribu d'événements $\mathfrak{F}_\tau$, tels que l'on puisse dire s'ils se sont réalisés ou non avant l'instant τ. On dira que cette tribu est engendrée par l'instant markovien τ. On a la définition formelle suivante.

DÉFINITION. *On appelle tribu $\mathfrak{F}_\tau$ engendrée par l'instant markovien τ la tribu de tous les événements B $\mathfrak{S}$-mesurables tels que pour tout $t \in T$ on a*

$$B \cap \{\tau \leqslant t\} \in \mathfrak{F}_t.$$

EXEMPLES. La quantité $\tau = t_0$, où t_0 ne dépend pas du hasard, $t_0 \in T$, est un instant markovien, et $\mathfrak{F}_\tau = \mathfrak{F}_{t_0}$.

En effet, on a ici $\{\tau \leqslant t\} = \varnothing \in \mathfrak{F}_t$ si $t < t_0$, et $\{\tau \leqslant t\} = \Omega \in \mathfrak{F}_t$ si $t_0 \leqslant t$, de sorte que $\tau = t_0$ est un instant markovien. D'autre part, $B \cap \{\tau \leqslant t\} \in \mathfrak{F}_t\ \forall t \in T$ si et seulement si $B \in \mathfrak{F}_{t_0}$

Signalons deux propriétés simples des instants markoviens (sur un même flot de tribus $\{\mathfrak{F}_t, t \in T\}$).

THÉORÈME 2.

a) *Si* $\tau_1 < \tau_2$, *alors* $\mathfrak{F}_{\tau_1} \subset \mathfrak{F}_{\tau_2}$.

b) *Si* τ_1 *et* τ_2 *sont deux instants markoviens, alors* $\tau_1 \vee \tau_2$ *et* $\tau_1 \wedge \tau_2$ *le sont également.*

En effet, supposons que $B \in \mathfrak{F}_{\tau_1}$. Il vient $B \cap \{\tau_1 \leqslant t\} \in \mathfrak{F}_t$. Donc

$$B \underset{t\in T}{\cap} \{\tau_2 \leqslant t\} = (B \cap \{\tau_1 \leqslant t\}) \cap \{\tau_2 \leqslant t\} \in \mathfrak{F}_t.$$

Et $B \in \mathfrak{F}_{\tau_2}$. Ceci prouve la proposition a). Soit d'autre part $\tau^* = \tau_1 \vee \tau_2$. Il vient alors $\{\tau^* \leqslant t\} = \{\tau_1 \leqslant t\} \cap \{\tau_2 \leqslant t\} \in \mathfrak{F}_t$ Si $\tau_* = \tau_1 \wedge \tau_2$, alors

$$\{\tau_* \leqslant t\} = \{\tau_1 \leqslant t\} \cup \{\tau_2 \leqslant t\} \in \mathfrak{F}_t \quad \forall t \in T. \blacksquare$$

Les instants markoviens permettent d'étudier les valeurs d'un processus aléatoire en ces instants. Supposons que T est un ensemble dénombrable, $\{\xi(t), t \in T\}$ un processus attaché au flot $\{\mathfrak{F}_t, t \in T\}$ et τ un instant markovien à valeurs dans T.

THÉORÈME 3. *La variable aléatoire* ξ_τ *est* $\mathfrak{F}_\tau$*-mesurable.*

En effet, pour tout a

$$\{\xi_\tau < a\} \cap \{\tau \leqslant t\} = \bigcup_{\substack{s \leqslant t \\ s\in T}} \{\xi(s) < a\} \cap \{\tau = s\}.$$

Comme pour $s < t$ $\{\xi(s) < a\} \in \mathfrak{F}_s \subset \mathfrak{F}_t$, $\{\tau = s\} \in \mathfrak{F}_s \subset \mathfrak{F}_t$, on a $\{\xi_\tau < a\} \cap \{\tau \leqslant t\} \in \mathfrak{F}_t$. $\blacksquare$

Soit une suite finie $\{\xi_t, t = 1, 2, \ldots, N\}$ qui est supposée être une $\mathfrak{F}_t$-submartingale. Considérons une suite monotone non décroissante d'instants $\mathfrak{F}_t$-markoviens $\tau_1 \leqslant \tau_2 \leqslant \ldots \leqslant \tau_p$ et posons $\eta_k = \xi_{\tau_k}$, $\mathfrak{F}_k^* = \mathfrak{F}_{\tau_k}$.

THÉORÈME 4. *La suite* $\{\eta_k, \mathfrak{F}_k^*, k = 1, 2, \ldots, p\}$ *est une submartingale.*

D é m o n s t r a t i o n. Nous savons que les η_k sont $\mathfrak{F}_k^*$-mesurables. Par ailleurs $\mathsf{E}\,|\eta_k| < \infty$. En effet,

$$\mathsf{E}\,|\eta_k| \leqslant \sum_{k=1}^{N} |\xi(t)| < \infty.$$

Reste à démontrer que $\mathsf{E}\{\eta_{k+1} \mid \mathfrak{F}_k^*\} \geqslant \eta_k$. On peut se borner au cas où $\tau_{k+1} - \tau_k$ prend les valeurs 0 et 1. En effet, si tel n'est pas le cas, introduisons une nouvelle suite d'instants markoviens $\sigma_1 = \tau_k$, $\sigma_2 = \tau_{k+1} \wedge (\sigma_1 + 1)$, $\ldots$, $\sigma_N = \tau_{k+1} = \tau_{k+1} \wedge (\sigma_{N-1} + 1)$. Alors $\{\sigma_k, k = 1, \ldots, N\}$ forment une suite monotone non décroissante d'instants markoviens tels que les différences $\sigma_{k+1} - \sigma_k$ ne prennent que les valeurs 0 ou 1. Si l'on prouve que la suite

$\{\xi_{\sigma_k}, k = 1, \ldots, N\}$ est submartingale, il s'en suivra que

$$\mathsf{E}\{\eta_{k+1} \mid \mathfrak{F}_k^*\} = \mathsf{E}\{\xi_{\sigma_N} \mid \mathfrak{F}_k^*\} \geqslant \xi_{\sigma_k} = \eta_k,$$

d'où le théorème.

Soit donc $\tau_{k+1} - \tau_k = 0$ ou 1. Pour tout $B \in \mathfrak{F}_k^*$ on a

$$\int_B (\eta_{k+1} - \eta_k)\, d\mathsf{P} = \int_{B \cap \{\tau_{k+1} - \tau_k = 1\}} (\eta_{k+1} - \eta_k)\, d\mathsf{P} =$$

$$= \sum_{r=1}^{N} \int_{B \cap \{\tau_k = r\} \cap \{\tau_{k+1} = r+1\}} (\xi_{r+1} - \xi_r)\, d\mathsf{P}.$$

Or

$$B \cap \{\tau_k = r\} \in \mathfrak{F}_r, \quad \{\tau_{k+1} = r+1\} = \{\tau_{k+1} > r\} \in \mathfrak{F}_r.$$

Donc

$$B \cap \{\tau_k = r\} \cap \{\tau_{k+1} = r+1\} \in \mathfrak{F}_r$$

et

$$\int_{B \cap \{\tau_k = r\} \cap \{\tau_{k+1} = r+1\}} (\xi_{r+1} - \xi_r)\, d\mathsf{P} \geqslant 0.$$

Par conséquent

$$\int_B (\eta_{k+1} - \eta_k)\, d\mathsf{P} \geqslant 0 \quad \forall B \in \mathfrak{F}_k^*.$$

Ce qui équivaut à

$$\mathsf{E}\{\eta_{k+1} \mid \mathfrak{F}_k^*\} \geqslant \eta_k. \ \blacksquare$$

COROLLAIRE. *Si* $\{\xi_t, \mathfrak{F}_t, t = 1, \ldots, N\}$ *est une martingale,* $\tau_1 \leqslant \ldots \leqslant \tau_p$ *une suite d'instants* $\mathfrak{F}_t$*-markoviens, alors* $\{\xi_{\tau_k}, \mathfrak{F}_{\tau_k}, k = 1, 2, \ldots, p\}$ *est une martingale.*

Quelques inégalités.

THÉORÈME 5. *Soit* $\{\xi_n, \mathfrak{F}_n, n = 1, \ldots, N\}$ *une submartingale. Alors*

$$\mathsf{P}\{\sup_{1 \leqslant n \leqslant N} \xi_n \geqslant C\} \leqslant \frac{\mathsf{E}\xi_N^+}{C}, \quad C \geqslant 0. \tag{6}$$

Si $\{\xi_n, \mathfrak{F}_n, n = 1, \ldots, N\}$ *est une supermartingale, alors*

$$\mathsf{P}\{\sup_{1 \leqslant n \leqslant N} \xi_n \geqslant C\} \leqslant \frac{2 \sup_{1 \leqslant n \leqslant N} \mathsf{E}|\xi_n|}{C}. \tag{7}$$

D é m o n s t r a t i o n.

a) Soit $\tau = k$ si $\xi_j \leqslant C$, $j = 1, \ldots, k-1$, et $\xi_k < C$. Si $\xi_k \leqslant C$ pour $k = 1, \ldots, N$, on pose $\tau = N$. τ est un instant marko-

vien et $\tau \leqslant N$. Donc le couple (ξ_τ, ξ_N) forme une submartingale. Par suite

$$CP\{\sup_{1\leqslant n\leqslant N} \xi_n \geqslant C\} \leqslant \int_{\{\xi_\tau \geqslant C\}} \xi_\tau \, dP \leqslant \int_{\{\xi_\tau \geqslant C\}} \xi_N \, dP \leqslant E\xi_N^+. \qquad (8)$$

b) Si la suite ξ_n est une supermartingale, alors

$$E\xi_0 \geqslant E\xi_\tau \geqslant CP\{\sup \xi_n \geqslant C\} + \int_{\{\sup \xi_n \leqslant C\}} \xi_N \, dP \geqslant CP\{\sup \xi_n \geqslant C\} - E\xi_N^-,$$

d'où il suit

$$CP\{\sup \xi_n \geqslant C\} \leqslant 2 \sup_{1\leqslant n\leqslant N} E|\xi_n|. \blacksquare$$

COROLLAIRE 1. *Si ξ_n est une martingale et $p \geqslant 1$, alors*

$$P\{\sup_{1\leqslant n\leqslant N} |\xi_k| \geqslant C\} \leqslant \frac{E|\xi_N|^p}{C^p}.$$

COROLLAIRE 2 (inégalité de Kolmogorov). *Si $\zeta_1, \zeta_2, \ldots, \zeta_N$ sont des variables aléatoires indépendantes*, $E\zeta_k = 0$, $k = 1, \ldots, N$, *alors*

$$P\{\sup_{1\leqslant k\leqslant N} |\zeta_1 + \zeta_2 + \ldots + \zeta_k| > C\} \leqslant \frac{1}{C^2} \sum_{k=1}^{N} \operatorname{Var} \zeta_k. \qquad (9)$$

L'inégalité (6) porte souvent le nom d'*inégalité de Kolmogorov pour les submartingales.*

Si maintenant ξ_n est une supermartingale quelconque, en vertu de l'inégalité (6) on a

$$CP\{\inf_n \xi_n \leqslant -C\} \leqslant E|\xi_N|,$$

ceci et l'inégalité (6) entraînent

$$P\{\sup_{1\leqslant n\leqslant N} |\xi_n| > C\} \leqslant \frac{3 \sup_{1\leqslant n\leqslant N} E|\xi_n|}{C}. \qquad (10)$$

Munissons l'espace des variables aléatoires définies sur $\{\Omega, \mathfrak{S}, P\}$ des normes attachées aux espaces $\mathcal{L}_p = \mathcal{L}_p\{\Omega, \mathfrak{S}, P\}$. Posons

$$\|\xi\|_p = \{E|\xi|^p\}^{1/p} = \left\{\int_\Omega |\xi(\omega)|^p \, dP\right\}^{1/p}.$$

THÉORÈME 6. *Si ξ_n est une submartingale, pour $p > 1$ on a*

$$\|\sup_{1\leqslant n\leqslant N} \xi_n^+\|_p \leqslant q \|\xi_N^+\|_p, \quad q = \frac{p}{p-1}, \qquad (11)$$

et

$$\|\sup_{1\leqslant n\leqslant N} \xi_n^+\|_1 \leqslant 2(1 + E\xi_N^+ (\operatorname{Log} \xi_N^+)^+). \qquad (12)$$

D é m o n s t r a t i o n. Soient $G(x)$, $x \geqslant 0$, une fonction continue monotone non décroissante, $G(0) = 0$, η une variable aléatoire non négative de fonction de répartition $F(x)$. Une intégration par parties donne

$$\mathsf{E}G(\eta) = \int_0^\infty G(x)\,dF(x) = \int_0^\infty (1 - F(x))\,dG(x). \tag{13}$$

Soit

$$\eta = \max_{1 \leqslant n \leqslant N} \xi_n^+.$$

En vertu de (8)

$$x\mathsf{P}\{\eta \geqslant x\} \leqslant \int_{\{\eta \geqslant x\}} \xi_N\,d\mathsf{P} \quad \forall x > 0.$$

Donc $1 - F(x) = \mathsf{P}\{\eta \geqslant x\} \leqslant \frac{1}{x} \int_{\{\eta \geqslant x\}} \xi_N\,d\mathsf{P}$ et

$$\mathsf{E}G(\eta) \leqslant \int_0^\infty \frac{1}{x}\Big(\int_{\{\eta \geqslant x\}} \xi_N\,d\mathsf{P}\Big)\,dG(x) \leqslant \int_0^\infty \mathsf{E}\chi(\eta - x)\,\xi_N^+ \frac{dG(x)}{x},$$

où $\chi(x) = 1$ pour $x \geqslant 0$ et $\chi(x) = 0$ pour $x < 0$.

$$\mathsf{E}G(\eta) \leqslant \mathsf{E}\xi_N^+ \int_0^\eta \frac{dG(x)}{x}.$$

En faisant $G(x) = |x|^p$ et en utilisant l'inégalité de Hölder on obtient

$$\|\eta\|_p^p \leqslant \frac{p}{p-1}\,\mathsf{E}\xi_N^+ \eta^{p-1} \leqslant q\,\|\xi_N^+\|_p\,\|\eta\|_p^{p-1},$$

d'où découle (11).

Traitons le cas $p = 1$. Utilisons de nouveau l'inégalité (8). On a

$$2C\mathsf{P}\{\eta \geqslant 2C\} \leqslant \int_{\{\eta \geqslant 2C\}} \xi_N\,d\mathsf{P} \leqslant \int_{\{\xi_N \geqslant C\}} \xi_N\,d\mathsf{P} + \int_{\{\xi_N < C,\ \eta \geqslant 2C\}} \xi_N\,d\mathsf{P} \leqslant \int_{\{\xi_N \geqslant C\}} \xi_N\,d\mathsf{P} + C\mathsf{P}\{\eta \geqslant 2C\},$$

de sorte que

$$C\mathsf{P}\{\eta \geqslant 2C\} \leqslant \int_{\{\xi_n \geqslant C\}} \xi_N\,d\mathsf{P}.$$

Appliquons l'égalité (13) à $\frac{\eta}{2}$. Comme dans ce cas

$$1-F(x)=\mathsf{P}\left\{\frac{\eta}{2}\geqslant x\right\}\leqslant\frac{1}{x}\int\limits_{\{\xi_N\geqslant x\}}\xi_N\,d\mathsf{P},$$

on a

$$\mathsf{E}G\left(\frac{\eta}{2}\right)\leqslant\int\limits_0^\infty \mathsf{E}\xi_N^+\chi(\xi_N^+-x)\frac{dG(x)}{x}.$$

Posons $G(x)=(x-1)^+$. Il vient

$$\mathsf{E}\left(\frac{\eta}{2}-1\right)^+\leqslant\mathsf{E}\xi_N^+\int\limits_1^{\xi_N^+\vee 1}\frac{dx}{x}=\mathsf{E}\xi_N^+(\operatorname{Log}\xi_N^+)^+.$$

Comme

$$\mathsf{E}\left(\frac{\eta}{2}-1\right)^+\geqslant\mathsf{E}\frac{\eta}{2}-1,$$

on obtient

$$\mathsf{E}\eta\leqslant 2\,(1+\mathsf{E}\xi_N^+(\operatorname{Log}\xi_N^+))^+. \blacksquare$$

Définition. *On appelle nombre d'intersections* $\nu\,[a,\,b[$ *de l'intervalle* $[a,\,b[$ $(a<b)$, *de haut en bas, par une famille de variables aléatoires* $\{\xi(t),\ t\in T\}$ *le suprémum de nombres s tels qu'existe une suite* $\{t_i,\ i=1,\ \ldots,\ 2s\}$, $t_i<t_{i+1}$, $t_i\in T$, *telle que*

$$\xi(t_1)>b,\ \xi(t_2)\leqslant a,\ \xi(t_3)>b,\ \ldots,\ \xi(t_{2s})\leqslant a.$$

Théorème 7 (inégalité de Doob). *Si* ξ_n, $n=1,\ \ldots,\ N$, *est une submartingale, alors*

$$\mathsf{E}\{\nu\,[a,\,b[\,|\,\mathfrak{F}_1\}\leqslant\frac{\mathsf{E}\{(\xi(N)-b)^+\,|\,\mathfrak{F}_1\}}{b-a}. \tag{14}$$

Démonstration. Sans restreindre la généralité on peut supposer que $a=0$ et $\xi_n\geqslant 0$. En effet, le cas général se ramène au précédent si l'on remplace ξ_n par $(\xi_n-a)^+$. Donnons-nous une suite d'instants markoviens $\tau_1,\ \tau_2,\ \ldots,\ \tau_N$ $(\tau_1\leqslant\tau_2\leqslant\ldots\leqslant\tau_N)$ de la façon suivante: τ_1 est le plus petit j tel que $\xi(j)>b$ si un tel j existe et $\tau_1=N$ si un tel j n'existe pas; τ_2 le plus petit j tel que $j>\tau_1$ et $\xi_j=0$ si un tel j existe, et $\tau_2=N$ si un tel j n'existe pas; τ_3 le plus petit j pour lequel $\xi_j>b$, $j>\tau_2$, $\tau_3=N$ si un tel j n'existe pas, et ainsi de suite. En vertu du théorème 1, on a

$$\mathsf{E}\,\{(\xi(\tau_2)-\xi(\tau_1))+(\xi(\tau_4)-\xi(\tau_3))+\ldots\,|\mathfrak{F}_1\}\geqslant 0.$$

D'un autre côté, la somme figurant sous le signe de l'espérance mathématique contient $\nu\,[a,\,b[$ termes inférieurs à $-b$ et, éventuellement,

au plus un terme non supérieur à $(\xi(N)-b)^+$. Donc

$$-b\mathsf{E}\{\nu]0,\ b]\mid\mathfrak{F}_1\}+\mathsf{E}\{(\xi(N)-b)^+\mid\mathfrak{F}_1\}\geqslant 0. \blacksquare$$

Les inégalités obtenues se généralisent facilement à des suites infinies. Ainsi, si $n=1, 2, \ldots$, l'inégalité (6) s'écrit

$$\mathsf{P}\{\sup_{1\leqslant n<\infty}\xi(n)\geqslant C\}\leqslant\frac{\sup\limits_{1\leqslant n<\infty}\mathsf{E}\xi^+(n)}{C}. \qquad (15)$$

La démonstration découle de

$$\sup_{1\leqslant n<\infty}\xi_n=\lim_{N\to\infty}\sup_{1\leqslant n<N}\xi_n,$$

donc

$$\mathsf{P}\{\sup_{1\leqslant n<\infty}\xi_n\geqslant C\}=\lim_{N\to\infty}\mathsf{P}\{\sup_{1\leqslant n\leqslant N}\xi_n>C\}.$$

De façon analogue, si $\nu_N\,]a,\ b]$ représente le nombre d'intersections de haut en bas de l'intervalle $]a,\ b]$ par la suite tronquée $\xi(1), \ldots, \xi(N)$ et $\nu_\infty\,]a,\ b]$ par toute la suite, alors du fait que $\nu_N\,]a,\ b]$ est monotone non décroissante lorsque N croît et $\nu_\infty\,]a,\ b]=$ $=\lim\limits_{N\to\infty}\nu_n\,]a,\ b]$, il suit

$$(b-a)\,\mathsf{E}\nu_\infty]\,a,\ b]\leqslant\sup_{1\leqslant n<\infty}\mathsf{E}\,(\xi_n-b)^+.$$

La situation est différente si la suite des valeurs de n ne possède pas de minimum mais un maximum. Soit $n=\ldots -k, -k+1, \ldots$ $\ldots, -1$ $(k>0)$. Les seconds membres des inégalités considérées atteignent leur maximum pour $n=-1$. Donc

$$\mathsf{P}\{\sup_{-\infty<n\leqslant -1}\xi_n\geqslant C\}\leqslant\frac{\mathsf{E}\xi^+_{-1}}{C}, \qquad (16)$$

$$\mathsf{E}\nu_\infty\,]\,a,\ b]\leqslant\frac{\mathsf{E}\,(\xi_{-1}-b)^+}{b-a}. \qquad (17)$$

Existence de la limite. Les théorèmes d'existence de la limite de la submartingale sont importants pour la suite.

On aura besoin du lemme suivant.

LEMME 1. *Pour qu'une suite numérique bornée $\{c_n,\ n=1, 2, \ldots\}$ soit convergente il faut et il suffit que pour tout couple de nombres rationnels a et b $(a<b)$ le nombre d'intersections de haut en bas de l'intervalle $]a,\ b]$ par cette suite soit fini.*

Soit $\nu\,]a,\ b]$ ce nombre. Si pour un couple quelconque $(a,\ b)$ on a $\nu\,]a,\ b]=\infty$, la suite c_n ne peut vérifier le critère de convergence. D'un autre côté, si c_n ne possède pas de limite, $\bar{c}=\overline{\lim}\ c_n>$

$> \overline{\lim}\, c_n = c$. Si a et b sont rationnels, $c < a < b < \bar{c}$, il existe une infinité de n_i et de n'_i tels que $c_{n_i} < a$, $c_{n'_i} > b$, c'est-à-dire $\nu\,]a, b] = \infty$. ∎.

THÉORÈME 8. *Soit* $\{\xi_n, n = 1, 2, \ldots\}$ *une submartingale. Alors*

a) *si* $\sup \mathsf{E}\xi_n^+ < \infty$, *alors* $\xi_\infty = \lim \xi_n$ *existe presque sûrement et* $\mathsf{E}\,|\,\xi_\infty\,| < \infty$;

b) *si une suite* $\{\xi_n, n = 1, 2, \ldots\}$ *est uniformément intégrable* (*ou équi-intégrable*), *alors* $\xi_\infty = \lim \xi_n$ *existe presque sûrement dans* $\mathcal{L}_1$ *et*

$$\xi_n \leqslant \mathsf{E}\,\{\xi_\infty \mid \mathfrak{F}_n\}, \quad n = 1, 2 \ldots$$

D é m o n s t r a t i o n. Prouvons tout d'abord que la suite $\{\xi_n\}$ est minorée. Remarquons à cet effet que

$$\{\xi_1 > b, \ \inf \xi_n = -\infty\} \subset \bigcap_{N=1}^{\infty} \{\nu\,]-N, b] > 0\}.$$

Par ailleurs

$$\mathsf{E}\nu\,]-N, b] \leqslant \frac{\mathsf{E}\,(\xi_n - b)^+}{b+N} \leqslant \frac{\mathsf{E}\xi_n^+ + b}{b+N} \to 0$$

lorsque $N \to \infty$. Donc

$$\lim \nu\,]-N, b] = 0$$

presque sûrement et comme $\nu\,]a, b]$ est un entier, il suit que

$$\nu\,]-N, b] = 0$$

presque sûrement pour N suffisamment grand. Donc pour tout b

$$\mathsf{P}\,\{\xi_1 > b, \ \inf \xi_n = -\infty\} = 0$$

et

$$\mathsf{P}\,\{\inf \xi_n = -\infty\} \leqslant \sum_{k=1}^{\infty} \mathsf{P}\,\{\xi_1 > -k, \ \inf \xi_n = -\infty\} = 0.$$

Nous avons démontré que la suite ξ_n est minorée. On établit de façon analogue qu'elle est majorée en se servant de l'inégalité de Kolmogorov

$$\mathsf{P}\,\{\sup \xi_n^+ > C\} \leqslant \frac{\sup \mathsf{E}\xi_n^+}{C}$$

qui entraîne

$$\mathsf{P}\,\{\sup \xi_n = +\infty\} = \lim_{C \to \infty} \mathsf{P}\,\{\sup \xi_n^+ > C\} = 0.$$

La suite $\{\xi_n, n = 1, 2, \ldots\}$ est donc presque sûrement bornée.

L'inégalité de Doob implique que $\mathsf{E}\nu\,]a, b] < \infty$ quels que soient a et b, de sorte que $\mathsf{P}\,(\Lambda_{ab}) = 0$, où $\Lambda_{ab} = \{\omega : \nu\,]a, b] = \infty\}$. Si donc $\Lambda = \bigcup_{a,\, b} \Lambda_{ab}$, où la sommation est étendue à tous les rationnels

a et b, $b > a$, alors $\mathsf{P}(\Lambda) = 0$. Le lemme 1 nous dit que $\lim \xi_n = \xi_\infty$ existe pour $\omega \overline{\in} \Lambda$. En outre

$$|\xi_n| = 2\xi_n^+ - \xi_n, \quad \mathsf{E}|\xi_n| = 2\mathsf{E}\xi_n^+ - \mathsf{E}\xi_n \leqslant 2 \sup \mathsf{E}\xi_n^+ - \mathsf{E}\xi_1 < \infty,$$

de sorte que

$$\mathsf{E}|\xi_\infty| = \mathsf{E} \lim |\xi_n| \leqslant \underline{\lim}\, \mathsf{E}|\xi_n| < \infty,$$

ce qui démontre la proposition a) du théorème.

Supposons maintenant que la suite $\{\xi_n, n = 1, 2, \ldots\}$ est équi-intégrable. Alors $\mathsf{E}\,|\,\xi_n\,| \leqslant C$ et $\lim \xi_n = \xi_\infty$ existe presque sûrement et $\mathsf{E}\,|\,\xi_\infty - \xi_n\,| \to 0$. D'où il suit que dans l'inégalité

$$\int_B \xi_m \, d\mathsf{P} \leqslant \int_B \xi_n \, d\mathsf{P}, \quad m < n, \quad B \in \mathfrak{F}_m,$$

on peut passer à la limite sous le signe de l'intégrale lorsque $n \to \infty$, de sorte que

$$\int_B \xi_m \, d\mathsf{P} \leqslant \underline{\lim} \int_B \xi_n \, d\mathsf{P} = \int_B \xi_\infty \, d\mathsf{P}. \ \blacksquare$$

COROLLAIRE. *Si* ξ_n, $n = 1, 2, \ldots$, *est une supermartingale non négative, la limite* $\xi_\infty = \lim \xi_n$ *existe presque sûrement.*

THÉORÈME 9. *Soient* $\{\xi_n, \mathfrak{F}_n, n = 1, 2, \ldots\}$ *une martingale et* $\mathfrak{F}_\infty$ *la plus petite tribu contenant toutes les tribus* $\mathfrak{F}_n$.

1. *Si* $\sup \mathsf{E}\,|\,\xi_n\,| < \infty$, *alors* $\lim \xi_n = \xi_\infty$ *existe presque sûrement.*

2. *Pour que la suite* ξ_n, $n = 1, 2, \ldots$, *soit équi-intégrable il est nécessaire et suffisant qu'existe une variable aléatoire intégrable* η *telle que*

$$\xi_n = \mathsf{E}\,\{\eta \mid \mathfrak{F}_n\}. \tag{18}$$

Si cette condition est réalisée, on peut poser $\eta = \xi_\infty$ *et* η *est définie de façon univoque* (mod P) *dans la classe des variables aléatoires* $\mathfrak{F}_\infty$*-mesurables.*

D é m o n s t r a t i o n. La proposition 1 découle du théorème 8. Si la suite ξ_n, $n = 1, 2, \ldots$, est équi-intégrable, d'après le théorème 8 on a presque sûrement $\xi_\infty = \lim \xi_n$ dans $\mathscr{L}_1$ et $\xi_n \leqslant \mathsf{E}\,\{\xi_\infty \mid \mathfrak{F}_n\}$. En appliquant cette inégalité à la suite $-\xi_n$ on obtient $\xi_n = \mathsf{E}\,\{\xi_\infty \mid \mathfrak{F}_n\}$, $n = 1, 2, \ldots$, $\mathsf{E}\,|\,\xi_\infty\,| < \infty$.

Soit η une variable aléatoire intégrable. Montrons que la suite $\xi_n = \mathsf{E}\,\{\eta \mid \mathfrak{F}_n\}$ est une martingale équi-intégrable. D'abord

$$\mathsf{P}\{|\xi_n| > C\} \leqslant \frac{\mathsf{E}|\xi_n|}{C} \leqslant \frac{\mathsf{E}|\eta|}{C} \to 0$$

uniformément en n lorsque $C \to \infty$. Ensuite

$$\int_{\{|\xi_n|>C\}} |\xi_n| \, d\mathsf{P} = \int_{\{\xi_n>C\}} \xi_n \, d\mathsf{P} - \int_{\{\xi_n<-C\}} \xi_n \, d\mathsf{P} =$$

$$= \int_{\{\xi_n>C\}} \eta \, d\mathsf{P} - \int_{\{\xi_n>-C\}} \eta \, d\mathsf{P} \leqslant \int_{\{|\xi_n|>C\}} |\eta| \, d\mathsf{P} \to 0$$

lorsque $C \to \infty$ uniformément en n, puisque $\mathsf{P}\{|\xi_n| > C\} \to 0$ lorsque $C \to \infty$ uniformément en n. Donc la suite ξ_n, $n = 1, 2, \ldots$, est équi-intégrable. Les égalités

$$\mathsf{E}\{\xi_{n+1} | \mathfrak{F}_n\} = \mathsf{E}\{\mathsf{E}\{\eta | \mathfrak{F}_{n+1}\} | \mathfrak{F}_n\} = \mathsf{E}\{\eta | \mathfrak{F}_n\} = \xi_n$$

entraînent que ξ_n est une $\mathfrak{F}_n$-martingale. Reste à montrer que η est définie de façon unique sur la classe des variables aléatoires $\mathfrak{F}_\infty$-mesurables.

Comme

$$\mathsf{E}\{\eta | \mathfrak{F}_n\} = \mathsf{E}\{\xi_\infty | \mathfrak{F}_n\}$$

pour tout n, on a

$$\int_A \eta \, d\mathsf{P} = \int_A \xi_\infty \, d\mathsf{P} \tag{19}$$

pour tout A appartenant à une tribu $\mathfrak{F}_n$. Or la classe d'ensembles A tels que (19) ait lieu est monotone et contient l'algèbre des ensembles $\bigcup_{n=1}^{\infty} \mathfrak{F}_n$. Elle contient donc la tribu $\sigma\{\bigcup_{n=1}^{\infty} \mathfrak{F}_n\}$. Donc (19) est réalisée pour tout $A \in \mathfrak{F}_\infty$. Et si η est une variable $\mathfrak{F}_\infty$-mesurable, alors $\eta = \xi_\infty \pmod{\mathsf{P}}$. ■

COROLLAIRE. *Soient* $\mathfrak{F}_n$ *un flot de tribus*, $\mathfrak{F}_\infty = \sigma\{\mathfrak{F}_n, n = 1, 2, \ldots\}$, η *une variable* $\mathfrak{F}_\infty$-*mesurable*, $A \in \mathfrak{F}_\infty$. *Alors dans* $\mathscr{L}_1$ *on a presque sûrement*

$$\lim \mathsf{E}\{\eta | \mathfrak{F}_n\} = \eta, \quad \lim \mathsf{P}\{A | \mathfrak{F}_n\} = \chi_A(\omega). \quad ■$$

Soit la submartingale

$$\ldots \xi_{-n}, \xi_{-n+1}, \ldots, \xi_{-1} \ (\mathfrak{F}_{-n} \subset \mathfrak{F}_{-n+1} \subset \ldots \subset \mathfrak{F}_{-1}). \tag{20}$$

Posons $\mathfrak{F}_{-\infty} = \bigcap_{n=1}^{\infty} \mathfrak{F}_{-n}$. Les résultats obtenus dans ce cas sont plus forts que les précédents.

THÉORÈME 10. *Si* $\inf \mathsf{E}\xi_n > -\infty$ *et la suite* (20) *est une submartingale, alors elle est équi-intégrable*, $\lim \xi_{-n} = \xi_{-\infty}$ *existe presque sûrement et dans* $\mathscr{L}_1$ *et*

$$\mathsf{E}\{\xi_{-n} | \mathfrak{F}_{-\infty}\} \geqslant \xi_{-\infty}. \tag{21}$$

D é m o n s t r a t i o n. Prouvons d'abord que la suite ξ_{-n}, $n = 1, 2, \ldots$, est équi-intégrable. Comme la suite $\mathsf{E}\xi_{-n}$ est monotone non croissante, alors existe $\lim \mathsf{E}\xi_{-n} = l > -\infty$. Pour tout $\varepsilon > 0$ on peut exhiber un n_0 tel que

$$0 \leqslant \mathsf{E}\,\xi_{-n_0} - \mathsf{E}\xi_{-n} < \varepsilon, \ n \geqslant n_0.$$

On remarquera que

$$\mathsf{E}\chi\,(|\,\xi_{-n}\,| > C)|\,\xi_{-n}\,| =$$
$$= -\mathsf{E}\xi_{-n} + \mathsf{E}\chi\,(\xi_{-n} \geqslant -C)\,\xi_{-n} + \mathsf{E}\chi\,(\xi_{-n} > C)\,\xi_{-n},$$

où $\chi\,(A)$ est l'indicateur de l'événement A. La définition de la submartingale entraîne que pour $n > n_0$

$$\mathsf{E}\chi\,(\xi_{-n} \geqslant -C)\,\xi_{-n} \leqslant \mathsf{E}\chi\,(\xi_{-n} \geqslant -C)\,\xi_{-n_0},$$
$$\mathsf{E}\chi\,(\xi_{-n} > C)\,\xi_{-n} \leqslant \mathsf{E}\chi\,(\xi_{-n} > C)\,\xi_{-n_0}.$$

Donc

$$\mathsf{E}\chi\,(|\,\xi_{-n}\,| > C)|\,\xi_{-n}\,| \leqslant$$
$$\leqslant \varepsilon - \mathsf{E}\xi_{-n_0} + \mathsf{E}\chi\,(\xi_{-n} \geqslant -C)\,\xi_{-n_0} + \mathsf{E}\chi\,(\xi_{-n} > C)\,\xi_{n_0} \leqslant$$
$$\leqslant \varepsilon + \mathsf{E}\chi\,(|\,\xi_{-n}\,| > C)|\,\xi_{-n_0}\,|.$$

D'un autre côté,

$$|\,\xi_{-n}\,| = 2\xi_{-n}^{+} - \xi_{-n}, \ \mathsf{E}|\,\xi_{-n}\,| \leqslant 2\mathsf{E}\xi_{-1}^{+} - l = k,$$

d'où en vertu de l'inégalité de Tchébychev

$$\mathsf{P}\,\{|\,\xi_{-n}\,| > C\} \leqslant \frac{\mathsf{E}\,|\,\xi_{-n}\,|}{C} \leqslant \frac{k}{C},$$

c'est-à-dire $\mathsf{P}\,\{|\,\xi_{-n}\,| > C\} \to 0$ lorsque $C \to \infty$ uniformément en n. Il existe donc un $C = C\,(\varepsilon)$ ne dépendant pas de n tel que

$$\mathsf{E}\chi\,(|\,\xi_{-n}\,| > C)|\,\xi_{-n_0}\,| < \varepsilon.$$

Donc

$$\mathsf{E}\chi\,(|\,\xi_{-n}\,| > C)|\,\xi_{-n}\,| \leqslant 2\varepsilon \ \forall\, n \geqslant n_0.$$

L'équi-intégrabilité de la suite ξ_{-n}, $n = 1, 2, \ldots$, est démontrée. En particulier, $\mathsf{E}\,|\,\xi_{-n}\,| \leqslant k$. L'existence presque sûre de la limite $\xi_{-\infty}$ de la suite ξ_{-n} découle de l'inégalité de Doob et se démontre exactement comme celle de la limite du théorème 7. L'équi-intégrabilité de la suite ξ_{-n} entraîne alors la convergence dans $\mathcal{L}_1$. De plus dans l'inégalité

$$\int_B \xi_{-n}\,d\mathsf{P} \leqslant \int_B \xi_{-k}\,d\mathsf{P}, \quad n > k, \quad B \in \mathfrak{F}_{-\infty},$$

on peut passer à la limite lorsque $n \to \infty$. On obtient

$$\int_B \xi_{-\infty}\,d\mathsf{P} \leqslant \int_B \xi_{-k}\,d\mathsf{P},$$

d'où dérive l'inégalité (21.) ■

COROLLAIRE. *Si* $\{\ldots, \xi_{-n}, \xi_{-n+1}, \ldots, \xi_{-1}\}$ *est une martingale, alors* $\xi_{-\infty} = \lim \xi_{-n}$ *existe presque sûrement et dans* $\mathcal{L}_1$, *la suite* ξ_{-n} *est équi-intégrable et* $\xi_{-\infty} = \mathsf{E}\{\xi_{-n} \mid \mathfrak{F}_{-\infty}\}$.

En effet, si ξ_{-n} est une martingale, alors ξ_{-n} et $-\xi_{-n}$ sont submartingales. Par ailleurs $\mathsf{E}\xi_{-n} = \text{const}$. En appliquant le théorème (10) aux suites ξ_{-n} et $-\xi_{-n}$ on obtient la proposition annoncée.

§ 2. Séries de variables indépendantes

On étudiera ici les critères de convergence presque sûre de séries à termes aléatoires indépendants.

Soit donnée la série

$$\xi_1 + \xi_2 + \ldots + \xi_n + \ldots \tag{1}$$

THÉORÈME 1. *Si existe une suite de nombres* $\varepsilon_n > 0$, $n = 1, 2, \ldots$, *telle que*

$$\sum_{n=1}^{\infty} \varepsilon_n < \infty, \quad \sum_{n=1}^{\infty} \mathsf{P}\{|\xi_n| > \varepsilon_n\} < \infty, \tag{2}$$

alors la série (1) *est presque sûrement absolument convergente.*

Démonstration. Soit $A_n = \{|\xi_n| > \varepsilon_n\}$. La convergence de la deuxième série de (2) et le théorème 8, § 1, chap. II, entraînent $\mathsf{P}(\overline{\lim}\, A_n) = 0$, c'est-à-dire seul un nombre fini d'événements A_n se réalisent presque sûrement. Il existe donc un $N = N(\omega)$ tel que pour $n > N(\omega)$ on a $|\xi_n| < \varepsilon_n$ et la série (1) est convergente. ■

Des résultats plus puissants ont lieu pour les semi-martingales. Posons

$$\zeta_n = \xi_1 + \xi_2 + \ldots + \xi_n, \quad \zeta_0 = 0, \quad \mathfrak{F}_n = \sigma\{\xi_1, \ldots, \xi_n\}.$$

THÉORÈME 2. *Supposons que* ξ_n *sont intégrables*, $n = 0, 1, \ldots$ *Alors:*

a) *si* $\mathsf{E}\{\xi_n \mid \mathfrak{F}_{n-1}\} \geqslant 0$ *et* $\sup \mathsf{E}\zeta_n^+ < \infty$, *la série* (1) *est presque sûrement convergente*;

b) *si* $\mathsf{E}\{\xi_n \mid \mathfrak{F}_{n-1}\} = 0$ *et pour* $p \geqslant 1$ $\sup_n \mathsf{E}|\zeta_n|^p < \infty$, *alors la série est presque sûrement convergente et convergente dans* $\mathcal{L}_p$.

La condition a) revient à supposer que $\{\zeta_n, \mathfrak{F}_n\}$ est une submartingale. La proposition correspondante est donc une conséquence du théorème de convergence de la submartingale. La condition b) signifie que $\{\zeta_n, \mathfrak{F}_n\}$ est une martingale. Donc $|\zeta_n|^p$ est une submartingale et

$$\mathsf{E} \sup_n |\zeta_n|^p \leqslant q^p \sup_n \mathsf{E}|\zeta_n|^p.$$

Donc $|\zeta_n|^p$ est équi-intégrable et ζ_n convergent presque sûrement et convergent dans $\mathscr{L}_p$ (théorème 16, § 1, chapitre II, théorème 9, § 1). ■

COROLLAIRE 1. *Si* $\mathsf{E}\{\xi_n \mid \mathfrak{F}_{n-1}\}=0$ *et* $\sum_{n=1}^{\infty} \mathsf{E}\xi_n^2 < \infty$, *alors la série* (1) *est convergente presque sûrement et convergente dans* $\mathscr{L}_2$.

La démonstration découle de ce que pour $k \neq n$

$$\mathsf{E}\xi_k\xi_n = \mathsf{E}\{\xi_k \mathsf{E}\{\xi_n \mid \mathfrak{F}_{n-1}\}\} = 0,$$

$$\mathsf{E}\zeta_n^2 = \mathsf{E}\left(\sum_{k=1}^{n} \xi_k\right)^2 = \sum_{k=1}^{n} \mathsf{E}\xi_k^2 + 2\sum_{j=2}^{n}\sum_{k<j} \mathsf{E}\xi_k\xi_j = \sum_{k=1}^{n} \mathsf{E}\xi_k^2,$$

et de la proposition b) du théorème. ■

Pour les séries à termes indépendants, ce résultat est connu sous le nom de théorème de Kolmogorov.

COROLLAIRE 2 (théorème de Kolmogorov). *Si* $\{\xi_n,\ n = 1, 2, \ldots\}$ *sont des variables aléatoires indépendantes,* $\mathsf{E}\xi_k = 0$ *et la série* $\sum_{k=1}^{\infty} \operatorname{Var} \xi_k < \infty$, *alors la série* (1) *est convergente presque sûrement.*

Cette assertion découle du corollaire 1 si $\mathfrak{F}_n$ est une tribu engendrée par les variables aléatoires $\xi_1, \xi_2, \ldots, \xi_n$ et de l'indépendance des variables aléatoires ξ_n :

$$\mathsf{E}\{\xi_n \mid \mathfrak{F}_{n-1}\} = \mathsf{E}\xi_n = 0.$$

Attardons-nous sur la convergence des séries à termes indépendants. Il suit de la loi de tout ou rien que ces séries sont convergentes ou bien avec la probabilité 1, ou bien avec la probabilité 0.

Nous aurons besoin d'une estimation de la répartition du maximum des sommes de termes indépendants.

THÉORÈME 3. *Si* $\{\xi_k,\ k = 1, 2, \ldots, n\}$ *sont indépendantes,* $\mathsf{E}\xi_k = 0$ *et* $|\xi_k| < c$ *presque sûrement, où c est une constante, alors*

$$\mathsf{P}\{\max_{1\leqslant k\leqslant n} |\zeta_k| \leqslant t\} \leqslant \frac{(c+t)^2}{\sum_{k=1}^{n} \sigma_k^2}, \tag{3}$$

où $\sigma_k^2 = \mathsf{E}\xi_k^2$.

Soient E_n les événements $\{\max_{0\leqslant k\leqslant n} |\zeta_k| \leqslant t\}$, $n = 1, 2, \ldots$ Ils forment une suite monotone décroissante. On a

$$\mathsf{E}\chi(E_n)\zeta_n^2=\sum_{k=1}^{n}\mathsf{E}\{\chi(E_k)\zeta_k^2-\chi(E_{k-1})\zeta_{k-1}^2\}=$$
$$=\sum_{k=1}^{n}\mathsf{E}\chi(E_{k-1})(\zeta_k^2-\zeta_{k-1}^2)-\sum_{k=1}^{n}\mathsf{E}\chi(E_{k-1}\diagdown E_k)\zeta_k^2. \quad (4)$$

Ensuite

$$\mathsf{E}\chi(E_{k-1}\diagdown E_k)\zeta_k^2=\mathsf{E}\chi(E_{k-1}\diagdown E_k)(\zeta_{k-1}+\xi_k)^2\leqslant(t+c)^2\mathsf{E}\chi(E_{k-1}\diagdown E_k),$$
$$\sum_{k=1}^{n}\mathsf{E}\chi(E_{k-1}\diagdown E_k)\zeta_k^2\leqslant(t+c)^2\sum_{k=1}^{n}\mathsf{E}\chi(E_{k-1}\diagdown E_k)=$$
$$=(t+c)^2[1-\mathsf{P}(E_n)]. \quad (5)$$

De plus

$$\mathsf{E}\chi(E_{k-1})(\zeta_k^2-\zeta_{k-1}^2)=\mathsf{E}\chi(E_{k-1})(2\zeta_{k-1}\xi_k+\xi_k^2)=$$
$$=2\mathsf{E}\chi(E_{k-1})\zeta_{k-1}\mathsf{E}\xi_k+\mathsf{E}\chi(E_{k-1})\mathsf{E}\xi_k^2=\sigma_k^2\mathsf{E}\chi(E_{k-1}). \quad (6)$$

Les relations (5) et (6) donnent

$$t^2\mathsf{P}(E_n)\geqslant\mathsf{E}\chi(E_n)\zeta_n^2\geqslant\sum_{k=1}^{n}\sigma_k^2\mathsf{E}\chi(E_{k-1})-(t+c)^2(1-\mathsf{P}(E_n))\geqslant$$
$$\geqslant\mathsf{P}(E_n)\left\{\sum_{k=1}^{n}\sigma_k^2+(t+c)^2\right\}-(t+c)^2,$$

ou

$$(t+c)^2\geqslant\mathsf{P}(E_n)\left\{\sum_{k=1}^{n}\sigma_k^2+c^2+2ct\right\}, \quad (7)$$

d'où suit (3). ■

Dans le cas général de séries à termes indépendants la convergence de la série (1) est tranchée par le théorème suivant.

THÉORÈME 4 (théorème de Kolmogorov des trois séries). *Pour qu'une série* (1) *de variables aléatoires indépendantes soit convergente presque sûrement il est nécessaire que pour tout* $c>0$ *et suffisant que pour un* $c>0$ *convergent les séries*

$$\sum_{n=1}^{\infty}\mathsf{P}\{|\xi_n>c\}, \quad (8)$$
$$\sum_{n=1}^{\infty}\mathsf{E}\xi_n', \quad (9)$$
$$\sum_{n=1}^{\infty}\mathrm{Var}\,\xi_n', \quad (10)$$

où $\xi_n'=\xi_n$ *pour* $|\xi_n|<c$ *et* $\xi_n'=0$ *pour* $|\xi_n|>c$.

Démonstration. *Condition suffisante.* Le corollaire 2 du théorème 1 dit que la série $\sum_{n=1}^{\infty}(\xi_n'-\mathsf{E}\xi_n')$ est convergente pres-

que sûrement, d'où, compte tenu de la convergence de la série (9), il suit que la série $\sum_{n=1}^{\infty} \xi_n'$ est convergente. La condition (8) et le théorème de Borel-Cantelli entraînent que seul un nombre fini de termes de la série $\sum_{n=1}^{\infty} (\xi_n - \xi_n')$ sont non nuls. Donc la série (1) est convergente presque sûrement.

Condition nécessaire. Supposons que la série (1) est convergente presque sûrement. Son terme général tendra alors vers 0 presque sûrement de sorte que seul un nombre fini de termes de cette série seront supérieurs à c en valeur absolue ($c > 0$). Donc la série $\sum_{n=1}^{\infty} \xi_n'$ est convergente presque sûrement. Désignons par $\{\eta_n\}$, $\eta = 1, 2, \ldots$, une suite de variables aléatoires indépendantes ne dépendant pas de la suite $\{\xi_n'\}$, $n = 1, 2, \ldots$, et admettant les mêmes répartitions que ξ_n'. Posons $\tilde{\xi}_n = \xi_n' - \eta_n$. La série $\sum_{n=1}^{n} \tilde{\xi}_n$ est alors convergente presque sûrement, $\mathsf{E}\tilde{\xi}_n = 0$, $|\tilde{\xi}_n| \leqslant 2c$, $\operatorname{Var} \tilde{\xi}_n = 2 \operatorname{Var} \xi_n'$. La convergence de la série $\sum_{n=1}^{\infty} \tilde{\xi}_n$ implique

$$\mathsf{P}\left\{\sup_{1 \leqslant n < \infty} \left|\sum_{k=1}^{n} \tilde{\xi}_k\right| < \infty\right\} = 1.$$

Donc pour un certain t

$$\mathsf{P}\left\{\sup_{1 \leqslant n < \infty} \left|\sum_{k=1}^{n} \tilde{\xi}_k\right| \leqslant t\right\} = a > 0.$$

De l'inégalité (3) on déduit pour tout n

$$2 \sum_{k=1}^{n} \operatorname{Var} \xi_k' = \sum_{k=1}^{n} \operatorname{Var} \tilde{\xi}_k \leqslant \frac{(2c+t)^2}{a},$$

ce qui prouve la convergence de la série (10). Le corollaire 2 du théorème 2 implique que la série $\sum_{n=1}^{\infty} (\xi_n' - \mathsf{E}\xi_n')$ est convergente presque sûrement. Ce qui à son tour entraîne la convergence de la série (9). Celle de la série (8) dérive du théorème de Borel-Cantelli, car si la série (1) converge, il existe presque sûrement seulement un nombre fini de termes de la série (1) tels que $|\xi_n| > c$. ■

COROLLAIRE. *Pour que la série* (1) *de variables aléatoires indépendantes non négatives soit convergente il est nécessaire que pour tout* $c > 0$

et suffisant que pour un $c > 0$ *les séries*

$$\sum_{n=1}^{\infty} \mathsf{P}\{\xi_n > c\}, \quad \sum_{n=1}^{\infty} \mathsf{E}\xi'_n$$

soient convergentes.

En effet, pour les ξ_n non négatifs on a $\mathsf{E}\xi_n'^2 \leqslant c\mathsf{E}\xi'_n$, donc la convergence de la série (9) entraîne celle de la série (10).

Il est possible grâce à une transformation simple de déduire à partir du théorème de la convergence presque sûre des séries un théorème du type loi forte des grands nombres, c'est-à-dire un théorème de convergence presque sûre de certaines moyennes de variables aléatoires. Voyons un exemple de cette nature.

LEMME 1. *Si la série* $\sum_{n=1}^{\infty} z_n$ *est convergente et* a_n *est une suite monotone croissante,* $a_n > 0$, $a_n \to \infty$, *alors*

$$\frac{1}{a_n} \sum_{k=1}^{n} a_k z_k \to 0.$$

Démonstration. Soit $S_0 = 0$, $S_n = \sum_{k=1}^{n} z_k$ et $|S_n| \leqslant c$, $n = 1, 2, \ldots$, où c est une constante. Posons

$$a_k - a_{k-1} = \Delta_k, \quad k = 1, 2, \ldots, a_0 = 0.$$

Il vient

$$\sum_{k=1}^{n} a_k z_k = \sum_{k=1}^{n} (\Delta_1 + \Delta_2 + \ldots + \Delta_k) z_k = \sum_{k=1}^{n} \Delta_k (S_n - S_{k-1}).$$

Donc

$$\left|\frac{1}{a_n} \sum_{k=1}^{n} a_k z_k\right| \leqslant \left|\frac{1}{a_n} \sum_{k=1}^{n_0} \Delta_k (S_n - S_{k-1})\right| + \sup_{n_0 \leqslant k \leqslant n} |S_n - S_{k-1}| \leqslant$$

$$\leqslant 2C \frac{a_{n_0}}{a_n} + \sup_{n_0 \leqslant k \leqslant n} |S_n - S_{k-1}| < \varepsilon$$

pour tout $\varepsilon > 0$ si n et n_0 sont choisis assez grands. ■

Ce lemme et le théorème de Kolmogorov (corollaire 2 du théorème 2) entraînent les assertions suivantes.

THÉORÈME 5. *Si* $\{\xi_n, n = 1, 2, \ldots\}$ *sont indépendantes et*

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \operatorname{Var} \xi_n < \infty,$$

alors

$$\lim \frac{1}{n} \sum_{k=1}^{n} (\xi_k - \mathsf{E}\xi_k) = 0$$

presque sûrement.

Des résultats plus puissants seront établis dans la suite comme conséquence des théorèmes ergodiques généraux pour les variables aléatoires ξ_n équi-réparties.

§ 3. Théorèmes ergodiques

Soit une suite aléatoire stationnaire $\{\xi(t),\ t \in T\}$, $T = \{t : t = 0, \pm 1, \ldots, \pm n, \ldots\}$ à valeurs dans un espace probabilisable $\{X, \mathfrak{B}\}$. La stationnarité de la suite (cf. chap. I, § 5) signifie que la répartition conjointe de la suite $\{\xi(t_1 + t),\ \xi(t_2 + t), \ldots, \xi(t_n + t)\}$ ne dépend pas de t quels que soient $n, t, t_1, \ldots, t_n$ $(t \in T,\ n > 0)$. Il revient au même de dire que pour toute fonction $f(x_1, \ldots, x_n)$, $x_k \in X$, bornée $\mathfrak{B}^n$ -mesurable la quantité $\mathsf{E}f(\xi(t_1 + t), \ldots, \xi(t_n + t))$ ne dépend pas de t quels que soient $n, t_1, \ldots, t_n$.

Soient X^T l'espace de toutes les suites $u = \{\ldots, x_{-n}, x_{-n+1}, \ldots, x_0, x_1, \ldots, x_n, \ldots\}$, $\mathfrak{C}$ la plus petite tribu contenant tous les ensembles cylindriques X^T, P_ξ une mesure induite sur $\mathfrak{C}$ par la suite $\{\xi(t),\ t \in T\}$. Donc l'espace probabilisé $\{X^T, \mathfrak{C}, \mathsf{P}\}$ représente naturellement le processus $\{\xi(t),\ t \in T\}$. Désignons par $\{X^T, \tilde{\mathfrak{C}}, \tilde{\mathsf{P}}_\xi\}$ l'espace muni de la mesure complétée. Définissons sur X^T une translation S dans le temps $S : u' = Su$ si $x'_n = x_{n+1}$, $n \in T$, où $u = \{x_n, n \in T\}$, $u' = \{x'_n, n \in T\}$. L'opération S possède une inverse S^{-1}: si $u'' = S^{-1}u$, $u'' = \{x''_n, n \in T\}$, alors $x''_n = x_{n-1}$. La stationnarité de la suite $\xi(t)$ signifie que pour tout ensemble cylindrique C

$$\mathsf{P}_\xi(C) = \mathsf{P}_\xi(SC). \tag{1}$$

Comme une mesure sur des ensembles cylindriques définit de façon unique une mesure sur $\mathfrak{C}$ et sur sa complétion $\tilde{\mathfrak{C}}_\xi$, l'égalité (1) est valable pour tout $A \in \tilde{\mathfrak{C}}_\xi$:

$$\mathsf{P}_\xi(A) = \mathsf{P}_\xi(SA), \quad A \in \tilde{\mathfrak{C}}_\xi. \tag{2}$$

D é f i n i t i o n. *Soient $\{U, \mathfrak{F}, \mu\}$ espace mesuré, S application mesurable de $\{U, \mathfrak{F}\}$ dans $\{U, \mathfrak{F}\}$. On dit que l'application S conserve la mesure si $\forall A \in \mathfrak{F}$*

$$\mu(S^{-1}A) = \mu(A),$$

où $S^{-1}A$ est l'antécédent complet de l'ensemble A.

Une application S est par définition *inversible* si existe une application mesurable S^{-1} telle que $SS^{-1} = S^{-1}S = I$, I étant l'identité. Dans ce cas l'application S^{-1} est dite l'*inverse de* S. La définition d'une suite stationnaire équivaut à la suivante: *une suite* $\{\xi(t),\ t \in T\}$ *est stationnaire si la mesure* P_ξ *est invariante par une translation dans le temps* S *dans* X^T.

L'étude des suites stationnaires est un cas particulier de l'étude des applications inversibles conservant la mesure (automorphismes) d'un espace mesuré.

Etudions le comportement asymptotique lorsque $n \to \infty$ de la moyenne

$$\frac{1}{n}\sum_{k=0}^{n-1} f(S^k u), \tag{3}$$

où S^k est la k-ième puissance de l'application S, $f(u)$ une fonction $\mathfrak{F}$-mesurable, $\{U, \mathfrak{F}, \mu\}$ un espace muni d'une mesure μ et $\mu(U) \leqslant \leqslant \infty$. Pour comprendre ce problème étudions le cas où $\{U, \mathfrak{F}, \mu\}$ est confondu avec $\{X^T, \widetilde{\mathfrak{C}}, \widetilde{\mathsf{P}}_\xi\}$ et S est une translation dans le temps. Soit $\xi_k = \xi(k, u) = x_k$, $f(u) = \chi_B(x_0)$, où $\chi_B(x)$ est l'indicateur d'un ensemble $B \in \mathfrak{B}$. On a

$$f(S^k u) = \chi_B(S^k u) = \chi_B(\xi(k))$$

et

$$\frac{1}{n}\sum_{k=0}^{n-1} f(S^k u) = \frac{\nu_n(B, u)}{n}, \tag{4}$$

où $\nu_n(B, u)$ est le nombre de termes de la suite $\xi(0), \xi(1), \ldots$ $\ldots, \xi(n-1)$ dont les valeurs appartiennent à B, c'est-à-dire que $\nu_n(B, u)$ est la fréquence d'accès des n premiers membres de la suite $\xi(t)$ $(t = 0, 1, \ldots, n-1)$ dans l'ensemble B. Donc le problème posé consiste en particulier à étudier le comportement de la fréquence du nombre de fois où les valeurs de la variable aléatoire $\xi(t)$ tombent dans un espace B. Montrons tout d'abord que la limite de (3) lorsque $n \to \infty$ existe presque sûrement. Cette proposition constitue le théorème classique de Birkhoff-Khintchine.

LEMME 1. *Si* S *préserve la mesure* μ, $D \in \mathfrak{F}$ *et* $f(u)$ *est une fonction* μ*-intégrable non négative* $\mathfrak{F}$*-mesurable, alors*

$$\int_{S^{-1}D} f(Su)\,\mu(du) = \int_D f(u)\,\mu(du). \tag{5}$$

Si l'on pose $f(u) = \chi_A(u)$, la formule (5) devient

$$\mu(S^{-1}(A \cap D)) = \mu(A \cap D),$$

égalité qui est valable quels que soient A et $D \in \mathfrak{F}$. D'où il suit que la formule (5) est valable pour toutes les fonctions μ-intégrables non négatives $\mathfrak{F}$-mesurables. ■

On se propose maintenant de démontrer un lemme de nature arithmétique. Soient $a_1, a_2, \ldots, a_n$ une suite de réels, p un nombre entier. On dit qu'un terme a_k de cette suite est *p-marqué* si l'une au moins des sommes

$$a_k,\ a_k + a_{k+1},\ \ldots,\ a_k + a_{k+1} + \ldots + a_{k+p-1}$$

n'est pas négative (a_k est 1-marqué si et seulement s'il n'est pas négatif).

LEMME 2. *La somme de tous les éléments p-marqués n'est pas négative.*

Soient a_{k_1} l'élément p-marqué de la suite de plus petit indice et $a_{k_1} + a_{k_1+1} + \ldots + a_{k_1+r}$ $(r \leqslant p-1)$ la somme non négative contenant le moins de termes. Si $h < r$, on a $a_{k_1} + a_{k_1+1} + \ldots + a_{k_1+h} < 0$, donc $a_{k_1+h+1} + \ldots + a_{k_1+r} \geqslant 0$, c'est-à-dire tous les termes de la suite $a_{k_1}, a_{k_1+1}, \ldots, a_{k_1+r}$ sont p-marqués et leur somme n'est pas négative. On peut reprendre ce raisonnement pour une suite de terme initial a_{k_1+r+1}. Donc toute la suite est décomposée en parties dont chacune s'achève par un groupe de termes p-marqués de somme non négative. L'ensemble des éléments p-marqués de toute la suite est confondu avec la somme des ensembles des éléments p-marqués de telles parties, d'où le lemme. ■

Le lemme suivant est la cheville ouvrière de la démonstration du théorème de Birkhoff-Khintchine.

LEMME 3. *Soient $f(u)$ fonction μ-intégrable, S application mesurable conservant la mesure μ de $\{U, \mathfrak{F}\}$ dans $\{U, \mathfrak{F}\}$ et*

$$E = \bigcup_{n=1}^{\infty} \Big\{u : \sum_{k=1}^{n} f(S^{k-1}u) \geqslant 0\Big\}.$$

Alors

$$\int_E f(u)\,\mu(du) \geqslant 0. \tag{6}$$

D é m o n s t r a t i o n. Considérons la suite $f(u), f(Su), \ldots$ $\ldots, f(S^{N+p-1}u)$ et désignons par $s(u)$ la somme de tous les éléments p-marqués de cette suite. En vertu du lemme 2 $s(u) \geqslant 0$. Soient $D_k = \{u : f(S^k u)$ est un élément p-marqué$\}$, $\chi_k(u)$ l'indicateur de l'ensemble D_k. On remarquera que

$$D_0 = \{u : \sup_{1 \leqslant k \leqslant p} f(S^{k-1}(u)) \geqslant 0\} \text{ et } D_k = S^{-1}D_{k-1} \text{ pour } k \leqslant N,$$

d'où

$$D_k = S^{-k}D_0 \quad (k \leqslant N),$$

et

$$0 \leqslant \int_U s(u)\,\mu(du) = \int_U \sum_{k=0}^{N+p-1} f(S^k u)\,\chi_k(u)\,\mu(du) =$$

$$= \sum_{k=0}^{N+p-1} \int_{D_k} f(S^k u)\,\mu(du).$$

Le lemme (1) donne

$$\int_{D_k} f(S^k u)\,\mu(du) = \int_{S^{-k}D_0} f(S^k u)\,\mu(du) = \int_{D_0} f(u)\,\mu(du), \quad k \leqslant N.$$

Donc

$$N \int_{D_0} f(u)\,\mu(du) + \sum_{k=N+1}^{N+p-1} \int_{D_k} f(S^k u)\,\mu(du) \geqslant 0. \tag{7}$$

Comme

$$\left| \int_{D_k} f(S^k u)\,\mu(du) \right| \leqslant \int_U |f(S^k u)|\,\mu(du) = \int_U |f(u)|\,\mu(du) < \infty,$$

en divisant l'inégalité (7) par N et en faisant tendre $N \to \infty$ on obtient

$$\int_{D_0} f(u)\,\mu(du) \geqslant 0. \tag{8}$$

Les ensembles $D_0 = D_0(p)$ $(p = 1, 2, \ldots)$ forment une suite monotone croissante et

$$\lim_{p\to\infty} D_0(p) = \bigcup_{p=1}^{\infty} D_0(p) = E.$$

Dans (8) en faisant tendre $p \to \infty$ on obtient (6). ■

Lemme 4 (théorème ergodique maximal). *Soient $f(u)$ fonction μ-intégrable, λ un réel et*

$$E_\lambda = \bigcup_{n=1}^{\infty} \left\{ u : \frac{1}{n} \sum_{k=1}^{n} f(S^{k-1}u) \geqslant \lambda \right\}.$$

Alors

$$\int_{E_\lambda} f(u)\,\mu(du) \geqslant \lambda\mu(E_\lambda). \tag{9}$$

On démontre ce lemme en appliquant le lemme (3) à la fonction $f(u) - \lambda$.

THÉORÈME 1 (théorème de Birkhoff-Khintchine). *Soient* $\{U, \mathfrak{F}, \mu\}$ *espace mesuré*, S *application mesurable de* $\{U, \mathfrak{F}\}$ *dans* $\{U, \mathfrak{F}\}$ *préservant la mesure* μ *et* $f(u)$ *fonction* μ*-intégrable. Alors* μ*-presque partout dans* U *existe la limite*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(S^k u) = f^*(u) \pmod{\mu}, \tag{10}$$

la fonction $f^*(u)$ *est* S*-invariante, c'est-à-dire*

$$f^*(Su) = f^*(u) \pmod{\mu}, \tag{11}$$

et est intégrable. Si $\mu(U) < \infty$, *alors*

$$\int_U f^*(u)\,\mu(du) = \int_U f(u)\,\mu(du). \tag{12}$$

D é m o n s t r a t i o n. Sans restreindre la généralité on peut admettre que la fonction $f(u)$ est finie et non négative. Posons

$$g^*(u) = \overline{\lim}\, \frac{1}{n} \sum_{k=0}^{n-1} f(S^k u), \qquad g_*(u) = \underline{\lim}\, \frac{1}{n} \sum_{k=0}^{n-1} f(S^k u).$$

Il faut établir que $g^*(u) = g_*(u) \pmod{\mu}$. Soit

$$K_{\alpha\beta} = \{u : g^*(u) > \beta,\ g_*(u) < \alpha\}, \qquad 0 \leqslant \alpha < \beta.$$

Il suffit de prouver que $\mu(K_{\alpha\beta}) = 0$. (En effet, $\{u : g^*(u) > g_*(u)\} = \bigcup_{\substack{\alpha<\beta \\ \alpha,\,\beta\in R}} K_{\alpha\beta}$, où R est l'ensemble des rationnels non négatifs.) On remarquera que

$$g^*(Su) = \overline{\lim}\, \left\{\frac{n+1}{n} \frac{1}{n+1} \sum_{k=0}^{n} f(S^k u) - \frac{f(u)}{n}\right\} = g^*(u)$$

et de façon analogue $g_*(Su) = g_*(u)$. Ceci signifie en particulier que $S^{-1}K_{\alpha\beta} = K_{\alpha\beta}$. On peut donc appliquer le lemme 4 à l'espace mesuré $\{K_{\alpha\beta}, \mathfrak{F} \cap K_{\alpha\beta}, \mu\}$. D'où il suit que

$$\int_{K_{\alpha\beta}} f(u)\,\mu(du) \geqslant \beta\mu(K_{\alpha\beta}). \tag{13}$$

En appliquant le lemme à la fonction $-f(u)$, on obtient

$$\int_{K_{\alpha\beta}} f(u)\,\mu(du) \leqslant \alpha\mu(K_{\alpha\beta}). \tag{14}$$

Comme $\beta > 0$, (13) implique que $\mu(K_{\alpha\beta}) < \infty$ et (14) n'a lieu que si $\mu(K_{\alpha\beta}) = 0$. D'où l'existence $\pmod{\mu}$ de la limite (10). Posons $f^*(u) = g^*(u)$. L'égalité (10) est alors réalisée et $f^*(u)$ est S-invariante presque partout dans U.

Pour prouver la formule (12) on posera

$$A_{kn}=\left\{u:\frac{k}{2^n}\leqslant f^*(u)<\frac{k+1}{2^n}\right\}.$$

On a

$$U=\bigcup_{k=-\infty}^{\infty}A_{kn},\quad S^{-1}A_{kn}=\left\{u:\frac{k}{2^n}\leqslant f^*(Su)<\frac{k+1}{2^n}\right\}=A_{kn}.$$

En appliquant le lemme 4 à l'ensemble A_{kn} on obtient pour tout $\varepsilon>0$

$$\int_{A_{kn}} f(u)\,\mu(du)\geqslant\left(\frac{k}{2^n}-\varepsilon\right)\mu(A_{kn}),$$

d'où lorsque $\varepsilon\to 0$

$$\int_{A_{kn}} f(u)\,\mu(du)\geqslant\frac{k}{2^n}\,\mu(A_{kn}).$$

De façon analogue

$$\int_{A_{kn}} f(u)\,\mu(du)\leqslant\frac{k+1}{2^n}\,\mu(A_{kn}),$$

d'où

$$\left|\int_{A_{kn}} f(u)\,\mu(du)-\int_{A_{kn}} f^*(u)\,\mu(du)\right|\leqslant\frac{1}{2^n}\,\mu(A_{kn}).$$

En sommant sur tous les k on obtient

$$\left|\int_U f(u)\,\mu(du)-\int_U f^*(u)\,\mu(du)\right|<\frac{1}{2^n}\,\mu(U)$$

La formule (12) découle de ce que n est arbitraire lorsque $\mu(U)\to\to\infty$. ∎

Quelques corollaires du théorème de Birkhoff-Khintchine.

Corollaire 1. *Supposons que* $\mu(U)<\infty$, $(u)\in\mathscr{L}_p\{U,\mathfrak{F},\mu\}$. *On a alors*

$$\int_U\left|\frac{1}{n}\sum_{k=0}^{n-1}f(S^ku)-f^*(u)\right|^p\mu(du)\to 0 \text{ lorsque } n\to\infty. \quad (15)$$

Pour démontrer ce corollaire considérons une fonction bornée quelconque $f_0(u)$ et soit $\|f(u)-f_0(u)\|_p=\delta$ où $\|f\|_p$ est la norme de

l'élément f dans $\mathscr{L}_p\{U, \mathfrak{F}, \mu\}$. Alors

$$\left\|\frac{1}{n}\sum_{k=1}^{n-1} f(S^k u) - f^*(u)\right\|_p \leqslant \left\|\frac{1}{n}\sum_{k=0}^{n-1}[f(S^k u) - f_0(S^k u)]\right\|_p + \\ + \left\|\frac{1}{n}\sum_{k=0}^{n-1} f_0(S^k u) - f_0^*(u)\right\|_p + \| f_0^*(u) - f^*(u)\|_p.$$

L'inégalité de Jensen et le lemme 1 donnent

$$\left\|\frac{1}{n}\sum_{k=0}^{n-1}[f(S^k u) - f_0(S^k u)]\right\|_p = \\ = \left\{\int_U \left|\frac{1}{n}\sum_{k=0}^{n-1}(f(S^k u) - f_0(S^k u))\right|^p \mu(du)\right\}^{1/p} \leqslant \\ \leqslant \left\{\int_U \frac{1}{n}\sum_{k=0}^{n-1}|f(S^k u) - f_0(S^k u)|^p \mu\, du)\right\}^{1/p} = \\ = \left\{\frac{1}{n}\sum_{k=0}^{n-1}\int_U |f(u) - f_0(u)|^p \mu(du)\right\}^{1/p} = \delta.$$

D'après le lemme de Fatou on a

$$\| f_0^*(u) - f^*(u)\|_p = \left\{\int_U \lim \left|\frac{1}{n}\sum_{k=0}^{n-1}[f(S^k u) - f_0(S^k u)]\right|^p \mu(du)\right\}^{1/p} \leqslant \\ \leqslant \underline{\lim} \left\|\frac{1}{n}\sum_{k=0}^{n-1}[f(S^k u) - f_0(S^k u)]\right\|_p \leqslant \delta.$$

La fonction $f_0(u)$ étant bornée, toutes ses moyennes sont majorées par la même constante. Donc dans l'expression

$$\left\|\frac{1}{n}\sum_{k=0}^{n-1} f_0(S^k u) - f_0^*(u)\right\|_p = \left\{\int_U \left|\frac{1}{n}\sum_{k=0}^{n-1} f_0(S^k u) - f_0^*(u)\right|^p \mu(du)\right\}^{1/p},$$

en vertu du théorème de Lebesgue on peut passer à la limite sous le le signe de l'intégrale lorsque $n \to \infty$. Par suite cette expression tend vers zéro et pour n assez grand devient inférieure à δ. Et

$$\left\|\frac{1}{n}\sum_{k=0}^{n-1} f(S^k u) - f^*(u)\right\|_p < 3\delta, \quad n \geqslant n_0 = n_0(\delta),$$

le nombre δ pouvant être pris aussi petit que l'on veut $(\delta > 0)$. D'où la formule (15). ∎

DÉFINITION 2. *Un ensemble* $A \in \mathfrak{F}$ *est S-invariant si* $\mu((S^{-1}A)\,\Delta A) = 0$.

Δ représente la différence symétrique d'ensembles.

On vérifie sans peine que la classe de tous les ensembles S-invariants est une tribu d'ensembles $\mathfrak{F}$-mesurables. D'autre part, si $g(u)$ est une fonction S-invariante, les ensembles $\{u : g(u) \geqslant c\}$ et $\{u : g(u) = c\}$ sont S-invariants. Par ailleurs, si A est S-invariant, $\chi_A(u)$ est une fonction S-invariante. Appelons $\mathfrak{I}$ la tribu des ensembles S-invariants. Soit $\mu(U) = 1$. On admettra que $\{U, \mathfrak{F}, \mu\}$ est un espace probabilisé et on désignera par E (espérance mathématique) l'intégration par rapport à la mesure μ.

COROLLAIRE 2. $f^*(u) = \mathsf{E}\{f(u) \mid \mathfrak{I}\}$ (mod μ).

De toute évidence $\mathsf{E}\{f(u) \mid \mathfrak{I}\}$ est une fonction S-invariante. Pour démontrer le corollaire 2 il suffit donc de vérifier que pour toute fonction S-invariante bornée on a

$$\mathsf{E}\, g(u)\,(f^*(u) - \mathsf{E}\{f(u) \mid \mathfrak{I}\}) = 0$$

ou encore que

$$\mathsf{E}\,(g(u) f^*(u) - g(u) f(u)) = 0;$$

or ceci découle de (12). Donc

$$(g(u) f(u))^* = \lim \frac{1}{n} \sum_{k=0}^{n-1} g(S^k u) f(S^k u) = g(u) f^*(u) \pmod{\mu}. \quad \blacksquare$$

Suites stationnaires ergodiques. Revenons aux suites stationnaires.

Soient $\{\xi(t), t \in T\}$ une suite stationnaire et $\{X^T, \mathfrak{C}, \mathsf{P}\}$ sa représentation naturelle.

COROLLAIRE 3. *Si f est une fonction mesurable dans $\{X^m, \mathfrak{B}^m\}$ et* $\mathsf{E}f(\xi(0), \xi(1), \ldots, \xi(m-1)) \neq \infty$, *alors presque sûrement*

$$\frac{1}{n} \sum_{k=0}^{n-1} f(\xi(k), \xi(k+1), \ldots, \xi(k+m-1)) \to$$

$$\to \mathsf{E}\{f(\xi(0), \xi(1), \ldots, \xi(m-1)) \mid \mathfrak{I}\} \text{ lorsque } n \to \infty,$$

où $\mathfrak{I}$ est la tribu des événements de $\mathfrak{F}$ invariants par une translation dans le temps.

Soient un événement $A \in \mathscr{C}$ et la suite d'événements déduits de A par une translation dans le temps: $A, S^{\pm 1}A, S^{\pm 2}A, \ldots$ Si χ_n est l'indicateur de l'événement $S^n A$, alors χ_n $(n = 0, \pm 1, \ldots)$ forment une suite stationnaire de variables aléatoires et $\frac{1}{n}\sum_{k=0}^{n-1} \chi_k$ est la

fréquence de réalisation de l'événement A, qui se calcule à l'aide d'une réalisation de la suite $\{\xi(t),\ t = 0, 1, 2, \ldots\}$:

$$\frac{1}{n}\sum_{k=0}^{n-1}\chi_k = \frac{\nu_n(A)}{n}.$$

Le théorème de Birkhoff-Khintchine entraîne l'existence presque sûre de

$$\lim_{n\to\infty}\frac{\nu_n(A)}{n} = \pi(A) = \mathsf{E}\{\chi_A \mid \mathfrak{I}\} \text{ et } \mathsf{E}\pi(A) = \mathsf{P}(A).$$

La quantité $\pi(A)$ est appelée *probabilité empirique* de l'événement A. C'est une variable aléatoire. Une question émerge tout naturellement: quand la probabilité empirique $\pi(A)$ ne dépend-elle pas du hasard et coïncide avec la probabilité $\mathsf{P}(A)$?

Les suites stationnaires possédant cette propriété s'appellent *ergodiques.*

La définition suivante est plus générale.

DÉFINITION. *Soient* $\{U, \mathfrak{F}, \mu\}$ *espace probabilisé,* S *application de* U *dans* U *préservant la mesure,* $\nu_n(A) = \nu_n(A, u)$ *le nombre de termes de la suite* $\{u, Su, \ldots, S^{n-1}u\}$, *contenus dans l'ensemble* A. *L'application* S *est ergodique si pour tout* $A \in \mathfrak{F}$

$$\lim_{n\to\infty}\frac{\nu_n(A, u)}{n} = \mu(A) \pmod{\mu}.$$

L'application S *est métrique transitive si tout ensemble* S*-invariant est de mesure* 1 *ou* 0.

THÉORÈME 2. *Pour qu'une application* S *définie sur l'espace probabilisé* $\{U, \mathfrak{F}, \mu\}$ *soit ergodique il est nécessaire et suffisant que soit réalisée l'une des deux conditions:*

a) S *est métrique transitive;*

b) *pour toute fonction* $f(u)$ μ*-intégrable* $\mathfrak{F}$*-mesurable la fonction*

$$f^*(u) = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} f(S^k u)$$

est presque sûrement constante.

D é m o n s t r a t i o n. Soit A un ensemble S-invariant et $0 < \mu(A) < 1$. Les différences symétriques des ensembles A, SA, $S^2A, \ldots$ sont de mesure nulle et $\nu_n(A) = n\chi_A(u) \pmod{\mu}$. Par suite $\lim\limits_{n\to\infty}\frac{\nu_n(A)}{n}$ ne peut être constante $(\mathrm{mod}\ \mu)$. Donc l'ergodicité entraîne la transitivité métrique. Supposons maintenant que S est métrique transitive. La fonction $f^*(u)$ étant S-invariante, la différence symétrique des ensembles

$$S^{-1}\{u : f^*(u) < x\} = \{u : f^*(Su) < x\} \text{ et } \{u : f^*(u) < x\}$$

est μ-négligeable. D'où il suit que $\mu\{u: f^*(u) < x\} = 0$ ou 1 pour tout x réel, c'est-à-dire $f^*(u) = \text{const} \pmod{\mu}$. Par conséquent a) implique b). Enfin la condition d'ergodicité est un cas particulier de la condition b), plus exactement, lorsque $f(u)$ est indicateur d'un événement. ■

Voyons quelques corollaires de l'ergodicité.

Soient $\{X, \tilde{\mathfrak{S}}, \mathsf{P}\}$ la représentation naturelle d'une suite stationnaire $\xi(n)$, S une translation dans le temps dans X^T, $\mathcal{L}_2 = \mathcal{L}_2\{X^T, \tilde{\mathfrak{S}}, \mathsf{P}\}$.

Le corollaire 1 du théorème 1 entraîne que

$$\lim_{n\to\infty} \int_{X^T} \frac{1}{n} \sum_{k=0}^{n-1} f(S^k u)\, g(u)\, \mathsf{P}(du) = \int_{X^T} f^*(u)\, g(u)\, \mathsf{P}(du) \tag{16}$$

pour des fonctions $f(u)$ et $g(u)$ de $\mathcal{L}_2$.

On dira qu'une suite $\{\xi(n), n = 0, \pm 1, \ldots\}$ est ergodique si l'application S l'est. Posons $g(u) = \eta$, $f(S^k u) = \zeta_k$ et supposons que la suite stationnaire initiale $\{\xi(n), n = 0, \pm 1, \ldots\}$ est ergodique. La relation (16) s'écrit

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \zeta_k \eta = \mathsf{E}\zeta_0 \mathsf{E}\eta. \tag{17}$$

Soient $g(u) = \chi_B(u)$, $f(u) = \chi_A(u)$, A et $B \in \tilde{\mathfrak{S}}$. De (17) il suit que

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathsf{P}(S^{-k}A \cap B) = \mathsf{P}(A)\,\mathsf{P}(B) \tag{18}$$

ou (si $\mathsf{P}(B) \neq 0$)

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathsf{P}(S^{-k}A \mid B) = \mathsf{P}(A), \tag{19}$$

où $\mathsf{P}(S^{-k}A \mid B)$ est la probabilité conditionnelle de l'événement $S^{-k}A$ par rapport à B.

LEMME 5. *L'égalité* (18) (*ou* (19)) *équivaut à l'ergodicité pour tous* $A, B \in \tilde{\mathfrak{S}}$.

Il suffit de montrer que (18) entraîne l'ergodicité. Soit C un événement S-invariant quelconque. Posons $A = B = C$ dans (18). Cette relation s'écrit alors :

$$\mathsf{P}(C) = \mathsf{P}^2(C),$$

d'où

$$\mathsf{P}(C) = 0 \text{ ou } 1$$

et le lemme découle du théorème 2. ■

L'égalité (19) admet l'interprétation probabiliste suivante. Soient A et B deux événements de $\tilde{\mathfrak{S}}$. Si l'on translate l'événement A indéfiniment dans le temps, les événements $S^{-n}A$ et B deviennent indépendants en moyenne quel que soit B.

La condition (19) est un cas particulier de la condition plus stricte:

$$\lim_{n\to\infty} \mathsf{P}(S^{-n}A|B) = \mathsf{P}(A), \tag{20}$$

appelée *condition de brassage*. La condition (20) est un cas particulier de l'égalité

$$\lim_{n\to\infty} \mathsf{E}\zeta_n\eta = \mathsf{E}\zeta_0\mathsf{E}\eta, \tag{21}$$

où $\zeta_n = f(S^n u)$, $\eta = g(u)$, $f(u)$ et $g(u)$ étant des fonctions quelconques de $\mathcal{L}_2$. D'autre part, (20) implique (21) pour des fonctions f et g étagées. En approchant $f(u)$ et $g(u)$ par des suites de fonctions simples $f_n(u)$ et $g_n(u)$ convergeant dans $\mathcal{L}_2$ respectivement vers $f(u)$ et $g(u)$, on s'assure que la condition de brassage équivaut à la condition (21). Il suffit de vérifier la condition (21) pour un ensemble de fonctions dont l'enveloppe linéaire est partout dense dans $\mathcal{L}_2$. Le mieux est de prendre pour une telle enveloppe les indicateurs d'ensembles cylindriques.

Soit la suite de variables aléatoires indépendantes équi-réparties $\{\xi_n, \eta = 0, \pm 1, \ldots\}$ et $\mathsf{E}\,|\,\xi_n\,| < \infty$. C'est une suite stationnaire. D'après le théorème de Birkhoff-Khintchine

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \xi_k = \xi^* \pmod{\mathsf{P}}, \quad \mathsf{E}\xi^* = \mathsf{E}\xi.$$

La variable aléatoire ξ^* ne dépend visiblement d'aucun nombre fini de variables $\xi_0, \xi_1, \ldots, \xi_p$. Donc ξ^* est mesurable par rapport à $\overline{\lim}\, \tilde{\sigma}\, \{\xi_k\}$ et en vertu de la loi de tout ou rien est constante, $\xi^* = c \pmod{\mathsf{P}}$ et de plus $c = \mathsf{E}\xi$. Nous avons donc le

THÉORÈME 3 (loi forte des grands nombres). *Si* $\{\xi_n, n = 0, \pm 1, \ldots\}$ *est une suite de variables indépendantes équi-réparties et* $\mathsf{E}\,|\,\xi_n\,| < \infty$, *alors*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \xi_k = \mathsf{E}\xi_0 \tag{22}$$

presque sûrement.

Le théorème démontré est une conséquence de l'ergodicité des variables aléatoires indépendantes équi-réparties. Mais on peut aller plus loin, plus exactement démontrer que l'opérateur de translation dans le temps dans X^T est de brassage. Ceci découle à son tour d'une proposition plus générale. Soient $\{\xi_n, n = 0, \pm 1, \ldots\}$ suite sta-

tionnaire d'éléments aléatoires dans $\{X, \mathfrak{S}\}$, $\mathfrak{F}_n$ tribu engendrée par les éléments aléatoires $\xi_n, \xi_{n+1}, \ldots, \mathfrak{F}_\infty = \bigcap_n \mathfrak{F}_n = \lim \mathfrak{F}_n$. On dira que la loi de tout ou rien s'applique à la suite $\{\xi_n, n = 0, \pm 1, \ldots\}$ si la tribu $\mathfrak{F}_\infty$ ne contient que des événements de probabilité 0 ou 1.

THÉORÈME 4. *Si une suite* $\{\xi_n, n = 0, \pm 1, \ldots\}$ *obéit à la loi de tout ou rien, toute translation dans le temps est de brassage.*

Posons $\zeta_{-n} = \mathsf{P}\{B \mid \mathfrak{F}_n\}$. La suite $\{\zeta_n, \mathfrak{F}_n, n = \ldots -k, -k+1, \ldots, 0\}$, $\mathfrak{F}_{-n} = \mathfrak{F}_n$, est une martingale et $\mathsf{P}\{B \mid \mathfrak{F}_\infty\}$ est son adhérence à gauche. La tribu $\mathfrak{F}_\infty$ étant triviale, il suit

$$\mathsf{P}\{B \mid \mathfrak{F}_\infty\} = \text{const} = \mathsf{P}(B) \pmod{\mathsf{P}}.$$

Le théorème de convergence des martingales (théorème 1, corollaire, § 1) implique $\lim \mathsf{P}\{B \mid \mathfrak{F}_n\} = \mathsf{P}(B)$ presque sûrement. Soit A un ensemble cylindrique sur les coordonnées $n = 0, 1, 2, \ldots$ Alors $S^{-n}A \in \mathfrak{F}_n$. Donc lorsque $n \to \infty$,

$$\mathsf{P}(B \cap S^{-n}A) = \int_{S^{-n}A} \mathsf{P}\{B \mid \mathfrak{F}_n\}\, \mathsf{P}(du) \to \mathsf{P}(B)\, \mathsf{P}(S^{-n}A) = \mathsf{P}(B)\, \mathsf{P}(A).$$

De toute évidence cette relation a lieu pour tout ensemble cylindrique A. D'où la relation (21) comme nous l'avons vu précédemment. ■

Comme autre exemple de processus vérifiant la condition de brassage citons une suite stationnaire gaussienne dont le coefficient de corrélation tend vers zéro. Soient $\{\xi_n, n = 0, \pm 1, \pm 2, \ldots\}$ une suite gaussienne stationnaire, $\mathsf{E}\xi_n = m$, $\mathsf{E}(\xi_n - m)(\xi_0 - m) = R_n$, $f(u) = f(x_0, x_1, \ldots, x_p)$, $g(u) = g(x_0, x_1, \ldots, x_p)$ des fonctions bornées suffisamment lisses de $p + 1$ variables dont les transformées de Fourier $f^*(\lambda_0, \ldots, \lambda_p)$, $g^*(\lambda_0, \ldots, \lambda_p)$ sont absolument intégrables. On a

$$\mathsf{E} f(\xi_n, \xi_{n+1}, \ldots, \xi_{n+p})\, g(\xi_0, \xi_1, \ldots, \xi_p) =$$

$$= \mathsf{E} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{i\left(\sum_{k=0}^{p} \lambda_k \xi_{n+k} + \sum_{k=0}^{p} \mu_k \xi_k\right)} \times$$

$$\times f^*(\lambda_0, \ldots, \lambda_p)\, g^*(\mu_0, \ldots, \mu_p)\, d\lambda_0 \ldots d\lambda_p\, d\mu_0 \ldots d\mu_p =$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left\{\sum_{k,r=0}^{p} R_{k-r}(\lambda_k\lambda_r + \mu_k\mu_r) + \sum_{k,r=0}^{p} R_{n+k-r}\lambda_k\mu_r\right\}} \times$$

$$\times f^*(\lambda_0, \ldots, \lambda_p)\, g^*(\mu_0, \ldots, \mu_p)\, d\lambda_0 \ldots d\lambda_p\, d\mu_0 \ldots d\mu_p.$$

Si $\lim\limits_{n\to\infty} R_n = 0$, en passant à la limite lorsque $n \to \infty$, on obtient

$$\lim_{n\to\infty} \mathsf{E}f(\xi_n, \xi_{n+1}, \ldots, \xi_{n+p})\, g(\xi_0, \xi_1, \ldots, \xi_p) =$$
$$= \mathsf{E}f(\xi_0, \xi_1, \ldots, \xi_p)\, \mathsf{E}g(\xi_0, \xi_1, \ldots, \xi_p). \quad (23)$$

La classe de fonctions f et g vérifiant la relation précédente étant partout dense dans $\mathscr{L}_2$, la relation (23) a lieu pour f et g quelconques de $\mathscr{L}_2$.

D'où le

Théorème 5. *Une suite gaussienne stationnaire dont le coefficient de corrélation $R_n \to 0$ lorsque $n \to \infty$ vérifie la condition de brassage.*

§ 4. Processus de renouvellement

En gros, le processus de renouvellement se laisse décrire de la façon suivante. On étudie les pannes d'un appareil. Dès qu'une panne se déclare, l'appareil est remplacé par un autre. On admettra que la durée de fonctionnement sans défaillance τ_n du n-ième appareil est une variable aléatoire et que les τ_n, $n = 1, 2, \ldots$, sont indépendantes et équi-réparties. Les instants de panne sont appelés *instants de renouvellement*. On convient que 0 est un instant de renouvellement. Posons

$$\xi_n = \tau_1 + \tau_2 + \ldots + \tau_n,\ n = 1, 2, \ldots,\ \xi_0 = 0.$$

La quantité ξ_n est l'instant du n-ième renouvellement. L'intervalle temporel $[0, \xi_n]$ comporte $n+1$ renouvellement (y compris le renouvellement à l'instant 0).

Soit $F(x) = \mathsf{P}\{\tau_n \leqslant x\}$ (dans ce paragraphe cette définition de la fonction de répartition d'une variable aléatoire est à préférer à la définition ordinaire $F(x) = \mathsf{P}\{\tau_n < x\}$. Il vient $F(x) = 0$ pour $x < 0$, $F(0) \geqslant 0$. On admettra que $F(0) < 1$. Alors $\mathsf{E}\tau_n > 0$.

Lemme 1. *$\xi_n \to \infty$ presque sûrement lorsque $n \to \infty$.*

En effet, $\mathsf{P}(\xi_n \leqslant c) = \mathsf{P}\left(\exp\left(-\sum_1^n \tau_k\right) \geqslant e^{-c}\right)$. L'inégalité de Tchébychev donne

$$\mathsf{P}(\xi_n \leqslant c) \leqslant e^c q^n, \text{ où } q = \mathsf{E}e^{-\tau_n}. \quad (1)$$

Donc

$$\mathsf{P}(\lim \xi_n \leqslant c) = \lim \mathsf{P}(\xi_n \leqslant c) = 0. \blacksquare$$

Le lemme 1 entraîne que pour tout $t \geqslant 0$ il existe un $\nu = \nu(t)$ tel que $\xi_{\nu-1} \leqslant t < \xi_\nu$. Appelons la quantité $\nu(t)$ *nombre de renouvellements* sur l'intervalle $[0, t]$. Comme $\{\nu(t) = n\} = \{\xi_{n-1} \geqslant t\} \setminus \{\xi_n \leqslant t\}$, il vient

$$\mathsf{P}\{\nu(t) = n\} = \mathsf{P}\{\xi_{n-1} \leqslant t\} - \mathsf{P}\{\xi_n \leqslant t\}.$$

ξ_n étant la somme de n variables indépendantes équi-réparties. Donc $\mathsf{P}(\xi_n \leqslant t) = F^{*(n)}(t)$, où $F^{*(n)}(t)$ est le n-uple produit de convolution de la fonction de répartition $F(x)$:

$$F^{*(n)}(t) = \int_{-\infty}^{\infty} F^{*(n-1)}(t-s)\,dF(s) = \int_0^t F^{*(n-1)}(t-s)\,dF(s),$$

où $F^{*(n)}(t) = 0$ pour $t < 0$, $F^{*(0)}(t) = I(t)$ ($I(t) = 1$ pour $t \geqslant 0$ et $I(t) = 0$ pour $t < 0$), $F^{*(1)} = F$.

Posons

$$H(t) = \begin{cases} \mathsf{E}\nu(t) & \text{pour } t \geqslant 0, \\ 0 & \text{pour } t < 0. \end{cases}$$

La fonction $H(t)$ jouera un rôle important dans la suite. On l'appelle *fonction de renouvellement*. C'est une fonction monotone non décroissante et continue à droite. Montrons qu'elle est finie quel que soit $t \geqslant 0$. Soit $\chi_n(t)$ l'indicateur de l'événement $\{\xi_n \leqslant t\}$. On a

$$H(t) = \mathsf{E}\nu(t) = \mathsf{E}\sum_{n=0}^{\infty} \chi_n(t) = \sum_{n=0}^{\infty} \mathsf{P}(\xi_n \leqslant t).$$

Donc

$$H(t) = \sum_{n=0}^{\infty} F^{*(n)}(t), \tag{2}$$

et en vertu de l'inégalité (1) la série est uniformément convergente sur tout intervalle fini $t \in [0, T]$. La convergence uniforme de la série (2) implique $\sum_{n=1}^{\infty} F^{*(n)}(t) = F * \left(\sum_{n=1}^{\infty} F^{*(n-1)}\right)(t) = F * H(t)$, où $F * G$ est le produit de convolution de F et G, nulles au-dessous de zéro:

$$F * G(x) = \int_{-\infty}^{\infty} G(x-s)\,dF(s) = \int_0^x G(x-s)\,dF(s).$$

La fonction $H(t)$ est donc solution de l'équation

$$H(t) = I(t) + \int_0^t H(t-s)\,dF(s). \tag{3}$$

Bien plus, si $z(t)$ est une fonction bornée mesurable sur tout compact $[0, T]$, $Z(t) = \int_0^t z(t-s)\,dH(s) = H * z(t)$, alors

$$Z(t) = z(t) + \int_0^t Z(t-s)\,dF(s), \quad t \geqslant 0. \tag{4}$$

Dans ce paragraphe l'intégration est étendue à l'intervalle fermé $[0, t]$. D'autre part, il convient de tenir compte de la valeur de la mesure engendrée par la fonction F et concentrée aux points 0 et t. L'équation (4) s'appelle *équation de renouvellement*. Elle possède une solution pour toute fonction $z(t)$ localement bornée mesurable (c'est-à-dire bornée sur tous les intervalles finis). On s'assure sans peine que dans la classe des fonctions bornées l'équation (4) admet une solution unique.

En effet, la différence $V(t)$ de deux solutions de (4) ($z(t)$ et $F(t)$ étant données) vérifie l'équation $V = F * V$. Donc $V = F^{*(n)} * V$ pour tout n. Or

$$\max_{t\in[0,\,T]} V(t) \leqslant \max_{t\in[0,\,T]} V(t)\, F^{*(n)}(T) \to 0 \text{ lorsque } n \to \infty,$$

donc

$$V(t) \equiv 0.$$

Etudions le comportement asymptotique de la fonction $H(t)$ lorsque $t \to \infty$.

LEMME 2. *La fonction $H(t)$ est semi-additive, c'est-à-dire*

$$H(t_1 + t_2) \leqslant H(t_1) + H(t_2).$$

Démonstration. On a

$$H(t_1+t_2) - H(t_1) = \sum_{n=0}^{\infty} \mathsf{P}(t_1 < \xi_n \leqslant t_1 + t_2) =$$

$$= \sum_{n=1}^{\infty} \sum_{k=1}^{n} \mathsf{P}(\xi_{k-1} \leqslant t_1 < \xi_k \leqslant t_1 + t_2,\ \xi_n \leqslant t_1 + t_2).$$

Comme $\xi_n - \xi_k$ ne dépend pas de ξ_k et ξ_{k-1}, il vient

$$\mathsf{P}(\xi_{k-1} \leqslant t_1 < \xi_k \leqslant t_1 + t_2,\ \xi_n \leqslant t_1 + t_2) =$$

$$= \mathsf{EP}(\xi_n - \xi_k \leqslant t_1 + t_2 - \xi_k \mid \xi_k)\, \chi(\xi_{k-1} \leqslant t_1 < \xi_k \leqslant t_1 + t_2),$$

où $\chi(A)$ désigne l'indicateur de l'événement A. D'autre part,

$$\sum_{n=k}^{\infty} \mathsf{P}(\xi_n - \xi_k \leqslant t_1 + t_2 - \xi_k \mid \xi_k) \leqslant H(t_1 + t_2 - \xi_k),$$

$$\sum_{k=1}^{\infty} \chi(\xi_{k-1} \leqslant t_1 < \xi_k \leqslant t_1 + t_2) \leqslant 1,$$

donc

$$H(t_1+t_2) - H(t_1) \leqslant$$

$$\leqslant \mathsf{E} \sum_{k=1}^{\infty} H(t_1 + t_2 - \xi_k)\, \chi(\xi_{k-1} \leqslant t_1 < \xi_k \leqslant t_1 + t_2) \leqslant H(t_2). \blacksquare$$

LEMME 3. *La limite* $l = \lim\limits_{t\to\infty} \frac{H(t)}{t}$ *existe pour toute fonction semi-additive localement bornée non négative* $H(t)$.

Démonstration. Soit $c > 0$, $t = kc + h$, $h \in [0, c[$ et k entier. On a alors

$$\frac{H(t)}{t} = \frac{H(kc+h)}{kc+h} \leqslant \frac{kH(c)+H(h)}{kc+h},$$

d'où

$$\overline{\lim_{t\to\infty}} \frac{H(t)}{t} \leqslant \frac{H(c)}{c} \quad \text{et} \quad \overline{\lim_{t\to\infty}} \frac{H(t)}{t} \leqslant \underline{\lim_{c\to\infty}} \frac{H(c)}{c}. \ \blacksquare$$

Calculons $\lim\limits_{t\to\infty} \frac{H(t)}{t}$.

THÉORÈME 1 (théorème élémentaire de la théorie du renouvellement).

$$\lim_{t\to\infty} \frac{H(t)}{t} = \frac{1}{\mathsf{E}\tau_1} \quad \left(\frac{1}{\mathsf{E}\tau_1} = 0 \ si \ \mathsf{E}\tau_1 = \infty\right). \tag{5}$$

Démonstration. Soit $\hat{H}(s)$ transformée de Laplace de la fonction $H(t)$:

$$\hat{H}(s) = \int_0^\infty e^{-st} H(t)\, dt.$$

On a

$$s^2 \hat{H}(s) = \int_0^\infty e^{-u} \left[\frac{H\left(\frac{u}{s}\right)}{\frac{u}{s}}\right] u\, du.$$

Comme

$$\int_0^\varepsilon e^{-u} \frac{H\left(\frac{u}{s}\right)}{\frac{u}{s}} u\, du = s^2 \int_0^{\varepsilon/s} e^{-st} H(t)\, dt \leqslant s\varepsilon H\left(\frac{\varepsilon}{s}\right) \to 0$$

avec ε et $H\left(\frac{u}{s}\right)\left(\frac{u}{s}\right)^{-1} \to l$ pour $s \to 0$ et $\forall u \geqslant 0$, il vient

$$\lim_{s\to 0} s^2 \hat{H}(s) = \int_0^\infty e^{-u} l u\, du = l.$$

Par ailleurs, en appliquant la transformation de Laplace à l'égalité (3) on obtient

$$\hat{H}(s) = \int_0^\infty e^{-st}\, dt + \int_0^\infty \int_0^t H(t-t')\, e^{-s(t-t')-st'}\, dF(t')\, dt =$$

$$=\frac{1}{s}+\int_0^\infty e^{-st'}\,dF(t')\int_0^\infty H(u)\,e^{-su}\,du=\frac{1}{s}+\hat{H}(s)\hat{F}(s),$$

où $\hat{F}(s)$ est la transformée de Laplace-Stieltjes de $F(t)$:

$$\hat{F}(s)=\int_0^\infty e^{-st}\,dF(t). \tag{6}$$

Donc

$$\hat{H}(s)=\frac{1}{s(1-\hat{F}(s))}. \tag{7}$$

Si $\mathsf{E}\tau_1$ est finie, alors

$$\lim_{s\downarrow 0}\frac{1-\hat{F}(s)}{s}=\int_0^\infty \frac{1-e^{-st}}{s}\,dF(t)\to\int_0^\infty t\,dF(t).$$

Dans ce cas

$$l=\lim_{s\to 0}s^2\hat{H}(s)=\lim_{s\to 0}\frac{s}{1-\hat{F}(s)}=\frac{1}{\mathsf{E}\tau_1}.$$

Si $\int_0^\infty t\,dF(t)=\infty$, il est aisé de voir que

$$\lim_{s\downarrow 0}\frac{1-\hat{F}(s)}{s}=\infty. \ \blacksquare$$

Ce résultat était prévisible: le nombre moyen de renouvellements dans l'unité de temps est égal à l'inverse du temps moyen de fonctionnement sans défaillance de l'appareil.

En théorie du renouvellement l'établissement de résultats plus précis est tributaire de la nature de la fonction de répartition $F(t)$.

Supposons que la durée de travail d'un appareil ne peut prendre que des valeurs de la forme $\tau = nh$, $n = 0, 1, \ldots$ On dit alors que les variables τ_k admettent une répartition réticulée. Sans restreindre la généralité on peut admettre que $h = 1$. Soit $p_n = \mathsf{P}(\tau_k = n)$. Posons

$$G(n) = 1 + \mathsf{P}(\tau_1 = n) + \ldots + \mathsf{P}(\tau_1 + \ldots + \tau_k = n) + \ldots$$

D'où il suit que $G(n) < \infty$ (si $\tau_k > 0$ presque sûrement, alors $G(n) \leqslant 2$). Soit d le plus grand commun diviseur (P.G.C.D.) des n tels que $p_n > 0$. Si $d = 1$, le processus de renouvellement sera dit *apériodique*, si $d > 1$, *périodique* et de *période de renouvellement*, d. Si le processus est apériodique, alors $G(n) > 0$ pour tous les n à partir d'un certain $n = n_0$. Si $d > 1$, alors $G(kd) > 0$ pour tous les

$k \geqslant k_0$ assez grands. Ces propositions découlent du lemme élémentaire de la théorie des nombres suivant.

LEMME 4. *Soit d le P.G.C.D. d'une suite de nombres entiers positifs $n_1, n_2, \ldots, n_s$. Il existe un nombre $m_0 > 0$ tel que pour tous les entiers $m \geqslant m_0$ l'égalité indéfinie*

$$md = \sum_{j=1}^{s} c_j n_j$$

admet une solution en nombres entiers non négatifs c_j.

D é m o n s t r a t i o n. Soit A l'ensemble de tous les nombres de la forme $x = \sum_1^s a_j n_j$, où a_j sont des entiers (positifs, négatifs ou nuls). Chaque x est divisible par d. Soit d_0 le plus petit nombre positif de A. Comme $x - kd_0 \in A$ pour tout k entier, quel que soit x, il existe un k tel que $x = kd_0$ (dans le cas contraire il existerait un k_1 tel que $x_1 = x - k_1 d_0$ satisfasse les inégalités $0 < x_1 < d_0$, ce qui contredit la définition de d_0). Donc d_0 est le P.G.C.D. des nombres de A. Soit maintenant $B = \{x : x = \sum_{j=1}^{s} b_j n_j\}$, où b_j sont des entiers non négatifs et $d_1 = \sum_{j=1}^{s} n_j$. On peut écrire: $d_0 = N_2 - N_1$, où $N_i \in B$. Soit c le plus grand des coefficients entiers en n_j intervenant dans N_2. Pour tout $m > 0$ entier posons $m = kd_1 + m_1$, où $0 \leqslant m_1 < d_1$. Alors $md_0 = kd_0 \cdot d_1 + m_1 d_0 \in B$ si $kd_0 > m_1 c_1$, ce qui est probablement réalisé pour $k > \frac{d_1 c}{d_0}$ ou pour $m > \frac{d_1^2 c}{d_0} + d_1$. ■

Considérons un processus apériodique de renouvellement et démontrons l'existence de $G_\infty = \lim_{n \to \infty} G(n)$.

LEMME 5. *Soient τ une variable aléatoire prenant la valeur n $(n = 0, \pm 1, \pm 2, \ldots)$ avec la probabilité p_n, $\varphi(u)$ la fonction caractéristique de τ. Si $d = 1$, alors $\varphi(u) \neq 1$ pour $|u| < 2\pi$, $u \neq 0$.*

D é m o n s t r a t i o n. On a

$$\varphi(u) = \mathsf{E} e^{iu\tau} = \sum_{-\infty}^{\infty} p_n e^{iun}.$$

Soit $\varphi(u_0) = 1$, $|u_0| < 2\pi$, $u_0 \neq 0$. On a

$$0 = 1 - \mathrm{Re}\, \varphi(u_0) = \sum_{-\infty}^{\infty} (1 - \cos nu_0)\, p_n.$$

Donc $\cos nu_0 = 1$ pour tous les n tels que $p_n > 0$ ou $nu_0 = 2\pi k$. Prenons une suite d'entiers $n_1, n_2, \ldots, n_s$ tels que $p_{n_r} > 0$ et que le P.G.C.D. soit égal à l'unité. On a alors $n_r u_0 = 2\pi k_r$ $(r = 1, 2, \ldots$

. . ., s). Par ailleurs, l'équation $\sum_{r=1}^{s} a_r n_r = 1$ admet une solution en nombres entiers a_r. Donc

$$u_0 = \sum_{r=0}^{s} a_r n_r u_0 = 2\pi \sum_{r=1}^{s} a_r k_r = 2\pi k_0,$$

où k_0 est un entier, ce qui contredit l'hypothèse $|u_0| < 2\pi$. ■

LEMME 6. *Si le renouvellement est apériodique*, $G_\infty = \lim_{n\to\infty} G(n)$ *existe.*

Démonstration. Posons

$$G(z, n) = \sum_{k=0}^{\infty} z^k p_n(k), \quad n \geqslant 0, \quad 0 \leqslant z \leqslant 1,$$

où $p_n(k) = \mathsf{P}\{\xi_k = n\}$, $\xi_k = \tau_1 + \ldots + \tau_k$. Le théorème d'Abel des séries entières donne

$$G(n) = \lim_{z \uparrow 1} G(z, n).$$

La fonction caractéristique de la variable aléatoire ξ_k étant égale à

$$[\varphi, (u)]^k = \sum_{n=1}^{\infty} p_n(k) e^{inu},$$

où $\varphi(u) = \mathsf{E} e^{iu\tau_k}$, on a

$$p_n(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-inu} [\varphi(u)]^k du.$$

Donc

$$G(z, n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{-inu} du}{1 - z\varphi(u)}, \quad n \geqslant 0.$$

L'intégrale du second membre est nulle pour $n < 0$. Donc

$$G(z, n) = \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{\cos nu \, du}{1 - z\varphi(u)}.$$

Posons

$$h(z, u) = \frac{1}{\pi} \operatorname{Re} (1 - z\varphi(u))^{-1}.$$

Comme $G(z, n)$ est une fonction réelle, on a

$$G(z, n) = \int_{-\pi}^{\pi} h(z, u) \cos nu \, du.$$

Le noyau $h(z, u)$ ($z \in [0, 1]$, $0 < |u| < 2\pi$) est positif et continu en vertu de l'apériodicité du renouvellement et du lemme 5. Donc pour $\varepsilon > 0$ quelconque

$$G(n) = \lim_{z \uparrow 1} \int_{-\varepsilon}^{\varepsilon} h(z, u) \cos nu \, du + \int_{\varepsilon \leqslant |u| \leqslant \pi} h(1, u) \cos nu \, du. \qquad (8)$$

En faisant $n = 0$ on voit qu'existe la limite

$$h_\varepsilon = \lim_{z \uparrow 1} \int_{-\varepsilon}^{\varepsilon} h(z, u) \, du,$$

et $h_\varepsilon \leqslant G(0)$. Comme h_ε décroît lorsque $\varepsilon \downarrow 0$, la limite $\lim_{\varepsilon \to 0} h_\varepsilon = h_0$ existe aussi. Donc existe

$$\lim_{\varepsilon \to 0} \lim_{z \uparrow 1} \int_{-\varepsilon}^{\varepsilon} h(z, u) \cos nu \, du = h.$$

Revenons à la formule (8). On constate que $h(1, u)$ est intégrable-Cauchy sur l'intervalle $]-\pi, \pi[$ et

$$G(n) = h + \int_{-\pi}^{\pi} h(1, u) \cos nu \, du.$$

$h(1, u)$ étant intégrable, le théorème de Riemann-Lebesgue donne

$$\lim_{n \to \infty} \int_{-\pi}^{\pi} h(1, u) \cos nu \, du = 0.$$

D'où l'existence de $\lim_{n \to \infty} G(n) = h$. ■

THÉORÈME 2. *Si le renouvellement est apériodique,*

$$\lim_{n \to \infty} G(n) = \frac{1}{m}, \quad m = \mathsf{E}\tau_h;$$

si $\mathsf{E}\tau_h = \infty$, $\lim_{n \to \infty} G(n) = 0$.

D é m o n s t r a t i o n. La limite $\lim_{n \to \infty} G(n) = h$ existant en vertu du lemme précédent, le théorème d'Abel des séries entières donne

$$h = \lim_{z \uparrow 1} \Big(1 + \sum_{n=1}^{\infty} z^n [G(n) - G(n-1)]\Big) =$$

$$= \lim_{z \uparrow 1} \sum_{n=0}^{\infty} z^n (1 - z) G(n) = \lim_{z \uparrow 1} (1 - z) \Phi(z),$$

où $\Phi(z)=\sum_{n=0}^{\infty} z^n G(n)$ est la génératrice de la suite $\{G(n),\ n=0,\ 1,\ \ldots\}$. L'indépendance et l'équi-répartition des τ_k entraînent que $G(n)$ est solution de l'équation

$$G(n)=\delta(n)+\sum_{k=1}^{n} G(n-k)\,p_k, \quad n\geqslant 0 \tag{9}$$

($\delta(n)=0$ pour $n>0$, $\delta(0)=1$). En multipliant par z^n et en sommant sur tous les $n\geqslant 0$, on obtient

$$\Phi(z)=1+F(z)\,\Phi(z), \quad |z|<1,$$

où $F(z)=\sum_{n=1}^{\infty} p_n z^n$. Donc

$$\Phi(z)=[1-F(z)]^{-1}$$

et

$$h=\lim_{z\uparrow 1}\left[\frac{1-F(z)}{1-z}\right]^{-1}.$$

Si $m=\infty$, pour tout $N>0$

$$\lim_{z\uparrow 1}\frac{1-F(z)}{1-z}\geqslant \lim_{z\uparrow 1}\sum_{n=1}^{N} p_n\frac{1-z^n}{1-z}=\sum_{n=1}^{N} p_n n,$$

d'où il suit que $h=0$. Si $m<\infty$, en tenant compte de ce que $\left|\frac{1-z^n}{1-z}\right|<n$ pour $|z|<1$, on obtient

$$\lim_{z\uparrow 1}\frac{1-F(z)}{1-z}=\lim_{z\uparrow 1}\sum_{1}^{\infty} p_n\frac{1-z^n}{1-z}=\sum_{n=1}^{\infty} p_n n=m. \blacksquare$$

COROLLAIRE. *Si le renouvellement est de période d,*

$$\lim_{n\to\infty} G(nd)=\frac{d}{m}, \quad m=\mathsf{E}\tau_k. \tag{10}$$

En effet, si le renouvellement est périodique et de période d, le nouveau renouvellement de durée $\tau'_n=\frac{\tau_n}{d}$ est apériodique. Si $G'(n)$ est sa fonction de renouvellement, alors $G'(n)=G(nd)$. Par ailleurs $\mathsf{E}\tau'_n=\frac{\mathsf{E}\tau_n}{d}=\frac{m}{d}$. Le théorème 2 entraîne (10). ■

Etudions la fonction de renouvellement dans le cas non réticulé. On fera cette étude à partir de l'équation de renouvellement (4). On aura besoin des lemmes suivants.

LEMME 7. *Soient* $\xi_1, \xi_2, \ldots$ *des variables aléatoires indépendantes équi-réparties,* $\zeta_n = f_n(\xi_1, \ldots, \xi_n)$, *où* $f_n(x_1, \ldots, x_n)$ *est une fonction symétrique, i.e.* $f(x_1, \ldots, x_n) = f(x_{i_1}, \ldots, x_{i_n})$ *pour toute permutation* $(i_1, \ldots, i_n)$ *des indices* $(1, 2, \ldots, n)$. *Si* $\zeta = \mathsf{P}\text{-}\lim \zeta_n$, *la variable ne dépend pas du hasard.*

D é m o n s t r a t i o n. On admet que $|f_n| \leqslant 1$ $\left(\text{dans le cas contraire on remplace } \zeta \text{ et } \zeta_n \text{ par } \frac{2}{\pi}\text{arc tg } \zeta \text{ et } \frac{2}{\pi}\text{ arc tg } \zeta_n\right)$. On a alors $\mathsf{E}\,|\zeta_{2n} - \zeta_n|^2 \to 0$. Or $\zeta_{2n} - \zeta_n$ admet la même répartition que $\zeta_{2n} - \zeta'_n$, où $\zeta'_n = f_n(\xi_{n+1}, \ldots, \xi_{2n})$. Donc $\lim\limits_{n\to\infty} \mathsf{E}(\zeta_{2n} - \zeta'_n)^2 = 0$ et par voie de conséquence $\lim\limits_{n\to\infty} \mathsf{E}(\zeta_n - \zeta'_n)^2 = 0$. L'indépendance de ζ_n et ζ'_n entraîne $\mathsf{E}(\zeta_n - \zeta'_n)^2 = (\mathsf{E}\zeta_n - \mathsf{E}\zeta'_n)^2 + \operatorname{Var}\zeta_n + \operatorname{Var}\zeta'_n \to 0$. D'où il suit $\operatorname{Var}\zeta = \lim \operatorname{Var}\zeta_n = 0$. ∎

LEMME 8. *Si une fonction* $\theta(t)$ *est continue, bornée et solution de l'équation*

$$\theta(t) = \int_0^\infty \theta(t-s)\, dF(s), \tag{11}$$

alors $\theta(t) \equiv \text{const.}$

D é m o n s t r a t i o n. L'équation (11) peut s'écrire

$$\theta(t) = \mathsf{E}\theta(t - \tau_k).$$

Soit $\mathfrak{F}_n$ la tribu engendrée par les variables aléatoires $\tau_1, \ldots, \tau_n$. On a alors

$$\mathsf{E}\{\theta(t-\xi_n) \mid \mathfrak{F}_{n-1}\} = \mathsf{E}\{\theta(t - \xi_{n-1} - \tau_n) \mid \mathfrak{F}_{n-1}\} =$$
$$= \mathsf{E}\theta(t - y - \tau_n)\,|_{y\ \xi_{n-1}} = \theta(t - \xi_{n-1}).$$

Donc la suite $\theta(t - \xi_n)$ est une martingale bornée. Par conséquent

$$\lim_{n\to\infty} \theta(t-\xi_n) = \hat{\theta}(t)$$

existe presque sûrement et dans $\mathscr{L}_1$. Comme $\hat{\theta}(t)$ est limite de fonctions symétriques dépendant de variables aléatoires équi-réparties, $\hat{\theta}(t)$ ne dépend pas du hasard (lemme 7). Donc

$$\hat{\theta}(t) = \lim \mathsf{E}\theta(t - \xi_n) = \theta(t)$$

et

$$\theta(t) = \lim_{n\to\infty} \theta(t-\xi_n) \text{ pour tout } t.$$

Vu que

$$\lim_{n\to\infty} \theta(t-\xi_n) = \lim \theta(t-\tau_1-(\tau_2+\ldots+\tau_n)) = \theta(t-\tau_1),$$

il suit que

$$\theta(t) = \theta(t-\tau_1) \text{ presque sûrement.}$$

La fonction $\theta(t)$ étant continue, on a

$$\mathsf{P}(\sup_t |\theta(t-\xi_1)-\theta(t)| > 0) = 0.$$

On peut donc trouver un ensemble A ($A \subset]-\infty, \infty[$) tel que

$$\theta(t) = \theta(t-x) \; \forall t \in]-\infty, \infty[, \; \forall x \in A.$$

La répartition de τ_1 étant non réticulée, A contient soit deux points incommensurables, soit des couples de points aussi proches que l'on veut. Donc la fonction continue $\theta(t)$ soit possède deux périodes incommensurables, soit des périodes aussi petites que l'on veut. Dans les deux cas cette fonction est constante. ∎

Le théorème suivant est connu sous le nom de *théorème fondamental de la théorie du renouvellement.*

THÉORÈME 3. *Si la répartition des variables τ_k est non réticulée, alors pour $c > 0$*

$$\lim_{t\to\infty} [H(t+c)-H(t)] = c/m, \quad m = \mathsf{E}\tau_k$$

(on convient que $c/m = 0$ lorsque $m = \infty$).

Démonstration. Soit $z(t)$ une fonction continue nulle en dehors d'un certain intervalle et $Z(t) = H * z(t)$. La fonction $Z(t)$ est bornée. En effet, si $z(t) = 0$ pour $|t| \geqslant A$ et $|z(t)| \leqslant C$, alors $|Z(t)| \leqslant C[H(t+A)-H(t-A)]$. D'un autre côté, la semi-additivité de la fonction $H(t)$ donne $H(t+A)-H(t-A) \leqslant \leqslant H(2A)$. Il est évident que la fonction $Z(t)$ est continue. Donc c'est l'unique solution de l'équation (4). Considérons la famille $\{H(s+t)-H(s), s \geqslant 0\}$ de fonctions monotones non décroissantes de t. Cette famille étant uniformément bornée ($H(s+t) - -H(s) \leqslant H(t)$) sur des intervalles quelconques $t \in [-A, A]$, il résulte que de toute suite de valeurs de s on peut extraire une suite partielle $s_k \to \infty$ telle que $H(s_k+t)-H(s_k)$ tendent vers une limite $v(t)$ sur un ensemble partout dense de valeurs de t. Or

$$Z(s_k+t) = \int_{-\infty}^{\infty} z(t-s)\, dH(s_k+s) \to \int_{-\infty}^{\infty} z(t-s)\, dv(s), \qquad (12)$$

puisque $z(t)$ est non nulle seulement sur un intervalle fini.

Posons

$$\theta(t) = \int_{-\infty}^{\infty} z(t-s)\, dv(s).$$

Dans l'équation

$$Z(s_k+t)=z(s_k+t)+\int_0^\infty Z(s_k+t-s)\,dF(s)$$

faisons tendre s_k vers l'infini. Comme $\lim z(s_k+t)=0$, on obtient

$$\theta(t)=\int_0^\infty \theta(t-s)\,dF(s).$$

Le lemme 8 donne $\theta(t)$ = const. Donc quel que soit $z(t)$ continue ($z(t)=0$ pour $|t|>A$, A quelconque), la fonction $\theta(t)= =\int_0^t z(t-s)\,dv(s)$ ne dépend pas de t. Par passage à la limite on conclut que $v(t_2)-v(t_1)$ dépend seulement de la différence t_2-t_1. Posons

$$v(t+s)-v(s)=h(t).$$

Il vient

$$h(t_1+t_2)=v(t_1+t_2+s)-v(t_1+s)+v(t_1+s)-v(s)= =h(t_2)+h(t_1).$$

La fonction $h(t)$ étant monotone non décroissante, les solutions de l'équation

$$h(t_1+t_2)=h(t_2)+h(t_1)\quad (t_1,\ t_2>0)$$

sont de la forme

$$h(t)=\alpha t.$$

Donc

$$v(t+s)-v(t)=\alpha s$$

et

$$Z(s_k+t)\to\alpha\int_{-\infty}^{\infty} z(t-s)\,ds,$$

la constante α ne dépendant ni du choix de la suite s_k, ni de la fonction $z(t)$. Dans la dernière relation s_k étant une suite partielle d'une suite quelconque $s_n'\to\infty$, on constate que la limite de $Z(t)$ lorsque $t\to\infty$ existe et

$$\lim_{t\to\infty} Z(t)=\alpha\int_{-\infty}^{\infty} z(s)\,ds. \qquad (13)$$

D'un autre côté les raisonnements précédents montrent que

$$\lim_{t\to\infty}[H(s+t)-H(t)]=\alpha s \qquad (14)$$

pour un ensemble dénombrable de valeurs de s partout dense dans $[-A, A]$. La continuité du second membre de (14) entraîne que l'égalité (14) est réalisée pour toutes les valeurs de s. On rappelle que (13) a été prouvée pour toutes les fonctions continues $z(t)$ non nulles sur un intervalle fini, et l'égalité (14) peut être considérée comme un cas particulier de (13) lorsque $z(s)$ est égal à l'indicateur de l'intervalle ouvert $]t, t+s]$. D'où il suit que l'égalité (13) a lieu, par exemple, pour la fonction $z(t) = 1 - F(t)$ avec $t \geqslant 0$, $z(t) = 0$ avec $t < 0$ si

$$m = \int_0^\infty (1 - F(t))\, dt = \mathrm{E}\tau_1 < \infty.$$

Dans ce cas

$$Z(t) = H * z(t) = \int_0^t (1 - F(t-s))\, dH(s) = 1$$

et l'égalité (13) entraîne $1 = \alpha m$, $\alpha = 1/m$. D'où le théorème pour $m < \infty$. Supposons maintenant que $m = \infty$. On a

$$1 = \int (1 - F(t-s))\, dH(s) \geqslant \lim \int_0^c [1 - F(s)]\, dH(t-s).$$

Lorsque $t \to \infty$, on obtient pour tout $c > 0$

$$1 \geqslant \alpha \int_0^c (1 - F(s))\, ds.$$

Or l'intégrale de droite tend vers l'infini avec c. Donc $\alpha = 0$. ■

Remarque. Pour démontrer le théorème on a appliqué la formule (13) à la fonction $z(t) = 1 - F(t)$, $t \geqslant 0$, alors que cette formule n'a été prouvée que pour des fonctions continues nulles en dehors d'un compact. Imposons la condition suivante.

Soient $h > 0$, c_{*n} et c_n^* respectivement le minimum et le maximum de la fonction $z(t)$ sur l'intervalle $(n-1)h \leqslant x \leqslant nh$.

Hypothèse A : les séries $\sigma_* = h \sum c_{*n}$, $\sigma^* = h \sum c_n^{*n}$ sont absolument convergentes et $\sigma^* - \sigma_* \to 0$ avec h.

Il est aisé de vérifier que si la fonction $z(t)$ satisfait l'hypothèse A, elle est justiciable de la formule (13). Prouvons-le. Supposons tout d'abord que la fonction $z(t)$ est en escalier, $z(t) = \sum a_n z_n(t)$, $z_n(t) = 1$ pour $h(n-1) \leqslant t \leqslant hn$ et $z_n(t) = 0$ pour les autres valeurs de t, $Z_n = H * z_n$. En vertu de (14) on a

$$Z_n(t) = H(t - (n-1)h) - H(t - nh) \to \alpha h,$$

$$Z_n(t) \leqslant M_h \quad \forall n,$$

où M_h est une constante. Soient $a_n \geqslant 0$, $\sum a_n < \infty$. On a

$$(Z = H * z) \quad \sum_1^n a_k Z_k(t) \leqslant Z(t) \leqslant \sum_1^n a_k Z_k(t) + M_h \sum_{n+1}^\infty a_k.$$

En faisant tendre $t \to \infty$ on obtient

$$\lim_{t\to\infty} Z(t) = \alpha h \sum_1^\infty a_k = \alpha \int_0^\infty z(x)\,dx.$$

Supposons maintenant que $z(t)$ est une fonction quelconque vérifiant l'hypothèse A. En appliquant le résultat précédent aux fonctions en escalier $\sum c_{*n} z_n(t)$ et $\sum c_n^* z_n(t)$, on trouve que tous les points limites de $Z(t)$ (lorsque $t \to \infty$) sont compris entre $\alpha\sigma_*$ et $\alpha\sigma^*$. Donc (13) est vraie pour des $z(t)$ quelconques vérifiant l'hypothèse A. De toute évidence $z(t) = 1 - F(t)$, $t \geqslant 0$, vérifie l'hypothèse A pourvu que $m = \mathsf{E}\tau_1 < \infty$. ∎

Introduisons maintenant deux caractéristiques importantes du processus de renouvellement. Si $\xi_{k-1} \leqslant t < \xi_k$, on posera

$$\gamma_t^+ = \xi_k - t, \quad \gamma_t^- = t - \xi_{k-1}.$$

Les processus γ_t^+ et γ_t^- sont linéaires par morceaux. Le processus γ_t^+ décroît sur tout intervalle de temps $[\xi_{k-1}, \xi_k[$ de τ_k à 0, le processus γ_t^- croît de 0 à τ_k. La quantité γ_t^+ indique combien de temps encore fonctionnera l'appareil si à l'instant t il est en marche, et la quantité γ_t^-, depuis combien de temps il fonctionne déjà. Dans certains problèmes on s'intéresse à la quantité $\gamma_t = \gamma_t^+ + \gamma_t^-$ qui est la durée totale de fonctionnement d'un appareil en marche à l'instant t. Comme γ_t est confondue avec un τ_k, il pourrait sembler que la répartition de γ_t coïncide avec celle de τ_k. En vérité ce n'est pas le cas. En effet, γ_t est confondue avec la valeur prise par τ_k à l'instant aléatoire $k = \nu(t) + 1$, $\gamma(t) = \tau_{\nu(t)+1}$, où $\nu(t)$ représente le nombre de renouvellements à l'instant t. Donc la répartition de γ_t peut ne pas être confondue avec celle de τ_k. Cette circonstance est connue sous le nom de *paradoxe de la théorie du renouvellement.*

Déterminons la répartition limite de γ_t, γ_t^+ et γ_t^-. Soit la fonction

$$V_t(x) = \mathsf{P}\{\gamma_t^+ > x\}.$$

L'événement $\{\gamma_t^- > y\} \cap \{\gamma_t^+ > x\}$ se réalise si et seulement si dans l'intervalle $[t - y, t + x]$ aucun renouvellement n'a lieu, c'est-à-dire lorsque $\gamma_{t-y}^+ > x + y$. Donc

$$\mathsf{P}\{\gamma_t^- > y, \gamma_t^+ > x\} = V_{t-y}(x + y).$$

La répartition conjointe de γ_t^+ et γ_t^- (et partant celle de $\gamma_t^+ + \gamma_t^-, \gamma_t^-$) s'exprime donc en fonction de $V_t(x)$. On remarquera que

$$\{\gamma_t^+ > x\} = \bigcup_{k=1}^{\infty} \{\xi_{k-1} \leqslant t\} \cap \{\xi_k > t+x\}.$$

Donc

$$V_t(x) = \sum_{k=0}^{\infty} \mathsf{P}(\xi_k \leqslant t, \xi_{k+1} > t+x).$$

En outre

$$\mathsf{P}(\xi_k \leqslant t, \xi_{k+1} > t+x) =$$

$$= \mathsf{E}\chi(\xi_k \leqslant t)\, \mathsf{P}(\tau_{k+1} > t+x-\xi_k \mid \xi_k\} = \int_0^t [1-F(t+x-s)]\, dF_k(s),$$

où $F_k(s)$ est la fonction de répartition de ξ_k, $F_k(s) = \mathsf{P}(\xi_k \leqslant s)$, $k \geqslant 1$ et $F_0(s)$ la répartition concentrée au point $s = 0$. Comme

$$\sum_{k=0}^{\infty} F_k(s) = H(s),$$

on a

$$V_t(x) = \int_0^1 (1-F(t+x-s))\, dH(s). \tag{15}$$

Théorème 4. *Si les variables τ_k possèdent une répartition non réticulée* et $\mathsf{E}\tau_k < \infty$, *alors*

$$\lim_{t\to\infty} V_t(x) = \frac{1}{\mathsf{E}\tau_1} \int_0^{\infty} [1-F(s+x)]\, ds. \tag{16}$$

La démonstration de ce théorème découle directement de la relation (13) si l'on pose $z(t) = 1 - F(t+x)$, $t \geqslant 0$, et l'on remarque que la formule (13) est applicable ici, puisque $z(t)$ vérifie l'hypothèse A.

Comme dans le cas réticulé on obtient la proposition suivante.

Théorème 5. *Si τ_k admettent une répartition réticulée,*

$$v_t(k) = \mathsf{P}(\gamma_t^+ = k),\quad p(k) = \mathsf{P}(\tau_1 = k),$$

alors

$$\lim_{t\to\infty} v_t(k) = \frac{1}{\mathsf{E}\tau_1} \sum_{j=0}^{\infty} p(k+j). \tag{17}$$

Etudions la répartition conjointe de γ_t^- et γ_t^+. On a

$$\{\gamma_t^- = j\} \cap \{\gamma_t^+ = k\} = \bigcup_{n=0}^{\infty} \{\zeta_n = t - j\} \cap \{\zeta_{n+1} = t + k\} =$$

$$= \bigcup_{n=0}^{\infty} \{\zeta_n = t - j\} \cap \{\tau_{n+1} = k + j\}.$$

Donc

$$\mathsf{P}\{\gamma_t^- = j,\ \gamma_t^+ = k\} = \sum_{n=1}^{\infty} p_{t-j}(n)\, p(k+j).$$

En vertu du théorème 2 on obtient

$$\lim_{t\to\infty} \mathsf{P}\{\gamma_t^- = j,\ \gamma_t^+ = k\} = \frac{1}{\mathsf{E}\tau_1} p(k+j).$$

D'où il suit

$$\lim_{t\to\infty} \mathsf{P}\{\gamma_t = l\} = \frac{l p(l)}{\mathsf{E}\tau_1}. \tag{18}$$

La formule (18) illustre le paradoxe de la théorie du renouvellement: d'une façon générale $\mathsf{P}\{\gamma_t = l\} \neq p(l)$.

§ 5. Chaînes de Markov

Le processus markovien au sens large a été défini au § 4, chapitre 1. Dans ce paragraphe on se propose d'étudier les chaînes de Markov ou processus markoviens à temps discret. Il s'agit d'un système stochastique dont les états sont décrits par les points d'un espace probabilisable $\{X, \mathfrak{B}\}$ appelé *espace des phases* du système. Le système est susceptible de changer d'états aux instants $t = 1, 2, \ldots$ La probabilité de passage $\mathsf{P}^{(m,n)}(x, B)$ est la probabilité conditionnelle que le système qui se trouve à l'instant m dans un état $x \in X$ passe à l'instant $n > m$ dans un des états de l'ensemble B, $B \in \mathfrak{B}$. On admet que $\mathsf{P}^{(m,n)}(x, B)$ à x fixe est une mesure sur $\mathfrak{B}$, $\mathsf{P}^{(m,n)}(x, X) = 1$, et à B fixe une fonction $\mathfrak{B}$-mesurable de x. La plus importante condition que l'on imposera à l'évolution de ce système est l'absence de post-action. Formellement cela revient à exiger que la probabilité de passage soit solution de l'équation de Kolmogorov-Chapman

$$\mathsf{P}^{(l,n)}(x, B) = \int_X \mathsf{P}^{(m,n)}(y, B)\, \mathsf{P}^{(l,m)}(x, dy), \quad l < m < n,$$

pour tous les $0 < l < m < n$, $x \in X$, $B \in \mathfrak{B}$.

Si le temps est discret, l'équation de Kolmogorov-Chapman montre que les probabilités de passage $\mathsf{P}^{(m,n)}(x, B)$ s'expriment par ré-

currence en fonction des probabilités de passage en une épreuve: $\mathsf{P}_n(x, B) \underset{\text{Déf}}{=} \mathsf{P}^{(n, n+1)}(x, B)$. En outre

$$\mathsf{P}^{(m, n+1}(x, B) = \int_X \mathsf{P}_n(y, B)\, \mathsf{P}^{(m, n)}(x\ dy).$$

Montrons que pour tout espace probabilisable $\{X, \mathfrak{B}\}$, toute mesure normée m sur $\mathfrak{B}$ et toute suite de noyaux stochastiques $\mathsf{P}_n(x, B)$, $n = 0, 1, \ldots$, on peut construire une chaîne de Markov pour laquelle les probabilités de passage en une épreuve sont égales à $\mathsf{P}_n(x, B)$ et la répartition initiale de l'état de la chaîne est donnée par la mesure m.

Mais d'abord nous allons munir un produit d'espaces d'une mesure à l'aide de noyaux stochastiques donnés.

Soit K une classe de fonctions réelles. On rappelle que K est un *cône* si elle contient les fonctions f et g et leurs combinaisons $af + bg$, où a et b sont des nombres non négatifs. La classe K est *monotone* si $f_n \in K$, $f_n \leqslant f_{n+1}$, $n = 1, 2, \ldots$, impliquent $\lim f_n \in K$.

LEMME 1. *Soient $\mathfrak{M}$ un semi-anneau de parties de X, $\mathfrak{A}$ la tribu engendrée par $\mathfrak{M}$, K une classe de fonctions définies sur X et possédant les propriétés suivantes*:

a) *c'est un cône et une classe monotone*;

b) $0 \leqslant f_1 \leqslant f_2$ *implique* $f_2 - f_1 \in K$;

c) $1 \in K$;

d) *elle contient les indicateurs des parties de $\mathfrak{M}$.*

Alors K contient toutes les fonctions $\mathfrak{A}$-mesurables non négatives.

Démonstration. Soit $\mathscr{L}$ la classe des parties de X dont les indicateurs appartiennent à K. On a

a) $X \in \mathscr{L}$;

b) si $A \in \mathscr{L}$, $B \in \mathscr{L}$ et $A \subset B$, alors $B \setminus A \in \mathscr{L}$ et $X \setminus A \in \mathscr{L}$;

c) si $A \in \mathscr{L}$ et $B \in \mathscr{L}$, alors $A \cap B \in \mathscr{L}$ et $A \cup B \in \mathscr{L}$. Donc $\mathscr{L}$ est une algèbre d'ensembles. La classe K étant monotone, il suit que $\mathscr{L}$ est une tribu. Donc K contient les indicateurs de tous les ensembles $\mathfrak{A}$-mesurables. Par suite elle contient toutes les fonctions $\mathfrak{A}$-mesurables simples et les limites de suites monotones non décroissantes de fonctions simples, c'est-à-dire toutes les fonctions $\mathfrak{A}$-mesurables non négatives.

LEMME 2. *Soient $f(x, y)$ une fonction non négative $\sigma\{\mathfrak{A} \times \mathfrak{B}\}$-mesurable et $\mathsf{P}(\cdot, \cdot)$ un noyau stochastique sur $\{X, \mathfrak{B}\}$. La fonction*

$$g(x) = \int_X f(x, y)\, \mathsf{P}(x, dy)$$

est $\mathfrak{A}$-mesurable.

Démonstration. A x fixe, la fonction $f(x, \cdot)$ est $\mathfrak{B}$-mesurable, de sorte que l'intégrale du second membre a un sens. La classe

K de fonctions non négatives $f(x, \cdot)$ obéissant au lemme est un cône et en vertu du théorème de Lebesgue est monotone. Par ailleurs, K contient la différence de deux de ses éléments et $f(x, y) \equiv 1 \in K$. Comme elle contient les indicateurs des ensembles de la forme $A \times B$, où $A \in \mathfrak{A}$, $B \in \mathfrak{B}$, elle contient toutes les fonctions non négatives $\sigma\{\mathfrak{A} \times \mathfrak{B}\}$-mesurables. ■

Les propositions suivantes généralisent le théorème de Fubini.

Théorème 1. *Soient* $\{X, \mathfrak{A}\}$, $\{Y, \mathfrak{B}\}$, $\{Z, \mathfrak{C}\}$ *des espaces probabilisables*, $Q_1(x, B)$, $Q_2(y, C)$ *des noyaux stochastiques sur* $\{X, \mathfrak{B}\}$ *et* $\{Y, \mathfrak{C}\}$ *respectivement. Il existe un noyau stochastique, unique de surcroît*, $Q_3(x, D)$ *sur* $\{X, \sigma\{\mathfrak{B} \times \mathfrak{C}\}\}$ *tel que*

$$Q_3(x, B \times C) = \int_B Q_1(x, dy)\, Q_2(y, C). \tag{1}$$

De plus, pour toute fonction $\sigma\{\mathfrak{B} \times \mathfrak{C}\}$*-mesurable non négative* $f(y, z)$ *on a*

$$\int_{Y\times Z} f(y, z)\, Q_3(x, dy \times dz) = \int_Y \Big(\int_Z f(y, z)\, Q_2(y, dz) \Big) Q_1(x, dy). \tag{2}$$

Pour démontrer la première partie du théorème il suffit de prouver que la formule (1) définit une prémesure semi-additive sur le semi-anneau des rectangles de $Y \times Z$. Soit $D_1 = B_1 \times C_1$, $D_2 = B_2 \times C_2$ et $D_2 \subset D_1$. On a alors $B_2 \subset B_1$, $C_2 \subset C_1$ et $D_1 = D_2 \cup D' \cup D''$, où $D' = B_2 \times (C_1 \setminus C_2)$, $D'' = (B_1 \setminus B_2) \times C_1$.
Les ensembles D_2, D' et D'' sont deux à deux disjoints. Si l'on applique la formule (1) successivement aux ensembles D_2, D' et D'', on obtient

$$Q_3(x, D_2) + Q_3(x, D') + Q_3(x, D'') =$$
$$= \int_{B_1} Q_1(x, dy)\, Q_2(y, C_2) + \int_{B_2} Q_1(x, dy)\, Q_2(y_1, C_1 \setminus C_2) +$$
$$+ \int_{B_1 \setminus B_2} Q_1(x, dy)\, Q_2(y, C) = \int_{B_1} Q_1(x, dy)\, Q_2(y, C) = Q_3(y, D_1).$$

Donc la fonction $Q_3(x, D)$ est additive sur les partitions considérées de l'ensemble D. En particulier, si $D_3 = D_1 \cup D_2$, où D_i sont des rectangles et $D_1 \cap D_2 = \varnothing$, alors

$$Q_3(x, D_1) + Q_3(x, D_2) = Q_3(x, D_3) \text{ et } Q_3(x, Y \times Z) = 1.$$

L'additivité de la fonction $Q_3(x, \cdot)$ sur le semi-anneau des rectangles se démontre sans peine dans le cas général par récurrence.

Soit $D = \bigcup_{k=1}^{n} D_k$, où D_k sont des rectangles deux à deux disjoints. Alors $D \setminus D_n = D' \cup D'' = \bigcup_{k=1}^{n-1} D_k$, où D' et D'' sont définis par les formules que nous venons tout juste d'utiliser. On a vu que

$$\mathrm{Q}_3(x, D) = \mathrm{Q}_3(x, D_n) + \mathrm{Q}_3(x, D') + \mathrm{Q}_3(x, D'').$$

En raisonnant par récurrence, on obtient

$$\mathrm{Q}_3(x, D') = \mathrm{Q}_3(x, D' \cap (\bigcup_{k=1}^{n-1} D_k)) = \sum_{k=1}^{n-1} \mathrm{Q}_3(x, D' \cap D_k).$$

L'expression de $\mathrm{Q}_3(x, D'')$ se déduit de façon analogue. Donc

$$\mathrm{Q}_3(x, D) = \mathrm{Q}_3(x, D_n) + \sum_{k=1}^{n-1} [\mathrm{Q}_3(x, D' \cap D_k) + \mathrm{Q}_3(x, D'' \cap D_k)].$$

D' et D'' étant des rectangles deux à deux disjoints dont la somme recouvre D_k, il suit que $D' \cap D_k$ et $D'' \cap D_k$ sont des rectangles aussi et $(D' \cap D_k) \cup (D'' \cap D_k) = D_k$. Donc

$$\mathrm{Q}_3(x, D' \cap D_k) + \mathrm{Q}_3(x, D'' \cap D_k) = \mathrm{Q}_3(x, D_k)$$

et

$$\mathrm{Q}_3(x, D) = \sum_{k=1}^{n} \mathrm{Q}_3(x, D_k).$$

D'où l'additivité de $\mathrm{Q}_3(x, \cdot)$. Prouvons maintenant la semi-additivité de $\mathrm{Q}_3(x, \cdot)$. Soit $D_0 \subseteq \bigcup_{1}^{\infty} D_k$, $D_k = B_k \times C_k$, $k = 0, 1, \ldots$ On a

$$\chi_{D_0}(y, z) \leqslant \sum_{k=1}^{\infty} \chi_{D_k}(y, z).$$

Comme

$$\chi_{D_k}(y, z) = \chi_{B_k}(y)\, \chi_{C_k}(z),$$

il vient

$$\chi_{B_0}(y)\, \chi_{C_0}(z) \leqslant \sum_{k=1}^{\infty} \chi_{B_k}(y)\, \chi_{C_k}(z).$$

En intégrant les deux membres par rapport à la mesure $\mathrm{Q}_2(y, \cdot)$ sur l'espace Z, on obtient

$$\chi_{B_0}(y)\, \mathrm{Q}_2(y, C_0) \leqslant \sum_{k=1}^{\infty} \chi_{B_k}(y)\, \mathrm{Q}_2(y, C_k).$$

L'intégration de cette expression par rapport à la mesure $Q_1(x, \cdot)$ nous conduit à

$$Q_3(x, D_0) \leqslant \sum_{k=1}^{\infty} Q_3(x, D_k),$$

inégalité qui traduit la semi-additivité dénombrable de $Q_3(x, D_k)$. D'où il suit que $Q_3(x, B \times C)$ admet un prolongement unique sur $\sigma\{\mathfrak{B} \times \mathfrak{C}\}$. Pour prouver la formule (2) on remarquera tout d'abord que d'après le lemme 2 l'intégrale intérieure du second membre de l'égalité (2) est une fonction $\mathfrak{B}$-mesurable, de sorte que l'intégrale double a un sens. La classe de fonctions f ($f \geqslant 0$) vérifiant la formule (2) satisfait les conditions a), b) et c) du lemme 1. D'après la formule (1) elle contient les indicateurs des rectangles. Donc elle contient toutes les fonctions $\sigma\{\mathfrak{B} \times \mathfrak{C}\}$-mesurables non négatives. ∎

On démontre de la même façon le

THÉORÈME 2. *Soient* $\{X, \mathfrak{A}\}$, $\{Y_1, \mathfrak{B}_1\}$, ..., $\{V_s, \mathfrak{B}_s\}$ *des espaces probabilisables, et* $Q_1(x, B^{(1)})$, $Q_2(y_1, B^{(2)})$, ..., $Q_s(y_{s-1}, B^{(s)})$ *des noyaux stochastiques,* $y_k \in Y_k$, $B^{(k)} \in \mathfrak{B}_k$ $(k = 1, \ldots, s)$. *Sur* $\{X, \mathfrak{D}\}$ *où* $\mathfrak{D} = \sigma\{\mathfrak{B}_1 \times \mathfrak{B}_2 \times \ldots \times \mathfrak{B}_s\}$ *il existe un noyau stochastique* $Q^{(1,s)}(x, D)$ *unique tel que*

$$Q^{(1,s)}(x, B^{(1)} \times \ldots \times B^{(s)}) = \int_{B^{(1)}} Q_1(x, dy_1) \int_{B^{(2)}} Q_2(y_1, dy_2) \ldots$$
$$\ldots \int_{B^{(s-1)}} Q_s(y_{s-1}, B^{(s)}) Q_{s-1}(y_{s-2}, dy_{s-1}). \quad (3)$$

De plus, pour toute fonction $f(y_1, y_2, \ldots, y_s)$ non négative *$\mathfrak{D}$-mesurable*

$$\int_{Y_1 \times \ldots \times Y_s} f(y_1, \ldots, y_s) Q^{(1,s)}(x, dy_1 \times \ldots \times dy_s) =$$
$$= \int_{Y_1} Q_1(x, dy_1) \ldots \int_{Y_s} f(y_1, \ldots, y_s) Q_s(y_{s-1}, dy_s). \quad (4)$$

REMARQUE. Les formules (2) et (4) ont été démontrées pour des fonctions non négatives. Elles sont de toute évidence valables pour des fonctions quelconques f si seulement l'une des fonctions f^+ ou f^- est intégrable. Il en sera de même dans les autres théorèmes où, pour abréger, l'on ne raisonnera que sur les fonctions non négatives.

Le noyau $Q^{(1,s)} = Q_1 \times Q_2 \times \ldots \times Q_s$ est appelé *produit direct, ou cartésien*, des noyaux $Q_1, Q_2, \ldots, Q_s$.

Si dans (3) on pose $B^{(1)} = Y_1, \ldots, B^{(s-1)} = Y_{s-1}$, on obtient un nouveau noyau stochastique sur $\{X, \mathfrak{B}_s\}$:

$$Q^{*(1,s)}(x, B^{(s)}) = Q^{(1,s)}(x, Y_1 \times Y_2 \times \ldots \times Y_{s-1} \times B^{(s)}) \quad (5)$$

que l'on appelle *produit de convolution* des noyaux $Q_1, Q_2, \ldots, Q_s$ et que l'on note

$$Q^{*(1,s)} = Q_1 * Q_2 * \ldots * Q_s.$$

En appliquant la formule (4) à la fonction $f(y_1, y_2, \ldots, y_s) = f(y_s) = \chi_{B^{(s)}}(y^{(s)})$ et en comparant avec (5) on obtient

$$\int\limits_{Y_1 \times Y_2 \times \ldots \times Y_s} f(y_s)\, Q^{(1,s)}(x, dy_1 \times dy_2 \times \ldots \times dy_s) = \int\limits_{Y_s} f(y_s)\, Q^{*(1,s)}(x, dy_s). \quad (6)$$

La classe des fonctions non négatives vérifiant la formule (6) satisfait les conditions du lemme 1. Donc la formule (6) a lieu pour toute fonction non négative $\mathfrak{B}_s$-mesurable. D'où il suit que pour toute fonction $f(y_{m_1}, y_{m_2}, \ldots, y_{m_r}, y_s)$ $(y_m \in Y_m,\ 0 \leqslant m_1 < m_2 < \ldots < m_r < s)$ de $r+1$ variables, non négative, $\sigma\{\mathfrak{B}_{m_1} \times \ldots \times \mathfrak{B}_{m_r} \times \mathfrak{B}_s\}$-mesurable, on a

$$\int\limits_{Y_1 \times Y_2 \times \ldots \times Y_{s-1} \times Y_s} f(y_{m_1}, y_{m_2}, \ldots, y_{m_r}, y_s) \times Q^{(1,s)}(x, dy_1 \times dy_2 \times \ldots \times dy_s) =$$

$$= \int\limits_{Y_{m_1}} Q^{*(1,m_1)}(x, dy_{m_1}) \int\limits_{Y_{m_2}} Q^{*(m_1+1,m_2)}(y_{m_1}, dy_{m_2}) \ldots \int\limits_{Y_s} f(y_{m_1}, \ldots, y_{m_r}, y_s)\, Q^{*(m_r+1,s)}(y_{m_r}, dy_s). \quad (7)$$

L'expression

$$Q^{*(1,s)} = Q^{*(1,m_1)} * Q^{*(m_1+1,m_2)} * \ldots * Q^{*(m_r+1,s)},$$

qui est un cas particulier de la formule (7), nous apprend que le produit de convolution est associatif.

Considérons des produits infinis de noyaux stochastiques. Soient $\{X_n, \mathfrak{B}_n\}$, $n = 0, 1, 2, \ldots$, une suite infinie d'espaces probabilisables, $\mathsf{P}_n(\cdot, \cdot)$, $n = 0, 1, 2, \ldots$, une suite de noyaux stochastiques définis sur $\{X_n, \mathfrak{B}_{n+1}\}$. Construisons d'après le théorème 2 les produits directs de noyaux

$$\mathsf{P}^{(0,n)} = \mathsf{P}_0 \times \mathsf{P}_1 \times \ldots \times \mathsf{P}_n,$$

$$\mathsf{P}^{(0,n)} = \mathsf{P}^{(0,n)}(x_0, D),\ x_0 \in X_0,\ D \in \mathfrak{C}_{n+1},$$

où $\mathfrak{C}_n$ est la plus petite tribu contenant les rectangles $B_1 \times B_2 \times \ldots \times B_n$ $(B_k \in \mathfrak{B}_k)$, $\mathfrak{C}_n = \sigma\{\mathfrak{B}_1 \times \mathfrak{B}_2 \times \ldots \times \mathfrak{B}_n\}$.

Soit l'espace $X^\infty = \prod_{n=1}^{\infty} X_n$ des suites $\omega = (x_1, x_2, \ldots, x_n, \ldots,)$, $x_n \in X_n$. Soit $\mathfrak{S}^0$ une algèbre d'ensembles cylindriques de X^∞. Définissons sur $\mathfrak{S}^0$ une famille de fonctions d'ensembles $\mathsf{P}^{(x_0)}$ dépendant d'un paramètre x_0 ($x_0 \in X_0$) comme suit: si C est l'ensemble cylindrique

$$C = \{\omega : (x_1, \ldots, x_n, \ldots) \in D\},\ D \in \mathfrak{S}_n,$$

on posera

$$\mathsf{P}^{(x_0)}(C) = \mathsf{P}^{(0,n)}(x_0, D).$$

Cette définition est univoque. En effet, si

$$C = \{\omega : (x_1, \ldots, x_n, \ldots) \in D'\},\ D' \in \mathfrak{S}_m,$$

et, par exemple $m > n$, alors

$$D' = D \times X_{n+1} \times \ldots \times X_m$$

et

$$\mathsf{P}^{(0,m)}(x_0, D') = \int_{X_1 \times \ldots \times X_m} \mathsf{P}_0(x_0, dx_1)\, \mathsf{P}_1(x_1, dx_2) \ldots$$
$$\ldots \mathsf{P}_{m-1}(x_{m-1}, dx_m)\, \chi_{D'}(x_1, \ldots, x_m),$$

où $\chi_{D'}(x_1, \ldots, x_m)$ est l'indicateur de D'. Vu que $\chi_{D'}(x_1, \ldots, x_m) = \chi_D(x_1, \ldots, x_n)$ et que $\mathsf{P}_{k-1}(x, X_k) = 1$, on déduit de la dernière expression

$$\mathsf{P}^{(0,m)}(x_0, D') = \mathsf{P}^{(0,n)}(x_0, D).$$

L'additivité de la fonction $\mathsf{P}^{(x_0)}$ sur $\mathfrak{S}^0$ est évidente.

Théorème 3. *Sur $\{X^\infty, \mathfrak{S}\}$, où $\mathfrak{S}$ est la tribu engendrée par les ensembles cylindriques de X^∞, il existe une famille unique de mesures $\mathsf{P}^{(x_0)}$ telle que*

$$\mathsf{P}_{(x_0)}\{\omega : x_k \in B_k,\ k = 1, \ldots, n\} = \int_{B_1} \mathsf{P}_0(x_0, dx_1) \int_{B_2} \mathsf{P}_1(x_1, dx_2) \ldots$$
$$\ldots \int_{B_{n-1}} \mathsf{P}_{n-2}(x_{n-2}, dx_{n-1})\, \mathsf{P}_{n-1}(x_{n-1}, B_n).$$

Démonstration. Il suffit de montrer que la mesure $\mathsf{P}^{(x_0)}$ dont est muni $\mathfrak{S}^0$ vérifie la condition de continuité: pour toute suite monotone décroissante d'ensembles cylindriques C_n tels que $\bigcap C_n = \varnothing$ on a $\mathsf{P}^{(x_0)}(C_n) \to 0$. Supposons l'inverse: $\mathsf{P}^{(x_0)}(C_n) \geqslant \varepsilon$ pour un x_0; soient D_n les bases des ensembles cylindriques C_n, $\chi(D_n; x_1, x_2, \ldots, x_{m_n}) = \chi(D_n)$ l'indicateur de D_n et D_n situées sur les

coordonnées $(1, 2, \ldots, m_n)$. Définissons une suite d'ensembles de $\mathfrak{B}$

$$B_n^{(1)} = \left\{ x_1 : \int\limits_{X^{(2, m_n)}} \chi(D_n ; x_1, x_2, \ldots, x_{m_n}) \times \right.$$

$$\left. \times \mathsf{P}^{(1, m_n)}(x_1, dx_2 \times \ldots \times dx_{m_n}) > \frac{\varepsilon}{2} \right\},$$

où $X^{(s, m)}$ représente le produit des espaces $X_s \times X_{s+1} \times \ldots \times X_m$.

La décroissance des C_n entraîne la décroissance monotone des $B_n^{(1)}$. D'autre part, si $\chi(B_n^{(1)})$ est l'indicateur de $B_n^{(1)}$ et $\bar{\chi}(B_n^{(1)}) = 1 - \chi(B_n^{(1)})$, alors

$$\varepsilon \leqslant \mathsf{P}^{(x_0)}(C_n) = \int\limits_{X_1} \int\limits_{X^{(2, m_n)}} (\chi(B_n^{(1)}) + \bar{\chi}(B_n^{(1)})) \times$$

$$\times \chi(D_n) \mathsf{P}_0(x_0, dx_1) \mathsf{P}^{(1, m_n)}(x_1, dx_2 \times \ldots \times dx_{m_n}) \leqslant$$

$$\leqslant \mathsf{P}_0(x_0, B_n^{(1)}) + \frac{\varepsilon}{2} \int\limits_{X_1} \bar{\chi}(B_n^{(1)}) \mathsf{P}_0(x_0, dx_1) \leqslant \mathsf{P}_0(x_0, B_n^{(1)}) + \frac{\varepsilon}{2}.$$

Donc $\mathsf{P}_0(x_0, B_n^{(1)}) > \frac{\varepsilon}{2}$. $\mathsf{P}_0(x_0, \cdot)$ étant une mesure, il s'ensuit que $\bigcap\limits_{n=1}^{\infty} B_n^{(1)} \neq \varnothing$. Soit $\bar{x}_1 \in B_n^{(1)}$, $n = 1, 2, \ldots$ Il vient

$$\int\limits_{X^{(2, m_n)}} \chi(D_n ; \bar{x}_1, x_2, \ldots, x_{m_n}) \mathsf{P}^{(1, m_n)}(\bar{x}_1, dx_2 \times \ldots \times dx_{m_n}) > \frac{\varepsilon}{2}.$$

Les raisonnements précédents valent pour le noyau $\mathsf{P}^{(2, m_n)}(x_2, dx_3 \times \ldots \times dx_{m_n})$ et la mesure $\mathsf{P}_1(\bar{x}_1, dx_2)$. Ils démontrent l'existence d'un point $\bar{x}_2$ tel que pour tout D_n l'on a

$$\int\limits_{X^{(3, m_n)}} \chi(D_n ; \bar{x}_1, \bar{x}_2, x_3, \ldots, x_{m_n}) \times$$

$$\times \mathsf{P}^{(2, m_n)}(\bar{x}_2, dx_3 \times \ldots \times dx_{m_n}) > \frac{\varepsilon}{4}.$$

Construisons la suite $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n, \ldots)$, $\bar{x}_n \in X_n$, telle que, quels que soient s et D_n,

$$\int\limits_{X^{(s+1, m_n)}} \chi(D_n ; \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_s, x_{s+1}, \ldots, x_{m_n}) \times$$

$$\times \mathsf{P}^{(s, m_n)}(\bar{x}_s, dx_{s+1} \times \ldots \times dx_{m_n}) \frac{\varepsilon}{2^s}.$$

Soit un ensemble C_k. Supposons que sa base D_k est située sur les coordonnées (1, 2, ..., s). La dernière inégalité nous dit que $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_s) \in D_k$ (sinon $\chi(D_k; \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_s, x_{s+1}, \ldots \ldots, x_{m_n}) \equiv 0$ pour tous les $(x_{s+1}, \ldots, x_m)$). Donc $(\bar{x}_1, \bar{x}_2, \ldots \ldots, \bar{x}_s, \ldots) \in C_k$ quel que soit C_k et $\bigcap_{k=1}^{\infty} C_k \neq \varnothing$, ce qui est contraire à l'hypothèse. ■

COROLLAIRE. *Soit donnée une suite dénombrable d'espaces probabilisés* $\{X_n, \mathfrak{B}_n, q_n\}$, $n = 1, 2, \ldots$ *Soient X l'espace de toutes les suites* $\omega = (x_1, x_2, \ldots, x_n, \ldots)$, $x_n \in X_n$, *et* $\mathfrak{C}$ *la tribu engendrée par les ensembles cylindriques* X^∞. *Sur* $\{X^\infty, \mathfrak{C}\}$ *il existe une mesure probabiliste stochastique unique* Q *telle que*

$$\mathsf{Q}\{\omega : x_k \in B_k, \ k = 1, 2, \ldots, n\} = \prod_{k=1}^{n} q_k(B_k), \ B_k \in \mathfrak{B}_k.$$

En d'autres termes, si est donnée une suite d'espaces probabilisés $\{X_n, \mathfrak{B}_n, q_n\}$, $n = 1, 2, \ldots$, il existe toujours un espace probabilisé $\{\Omega, \mathfrak{S}, \mathsf{Q}\}$ et une suite d'applications f_n de Ω dans X_n telles que les éléments aléatoires $\xi_n = f_n(\omega)$ admettent les répartitions q_n données sur $\mathfrak{B}_n$, et $\{\xi_n, n = 1, 2, \ldots\}$ sont indépendants dans leur ensemble.

REMARQUE. Contrairement au théorème de Kolmogorov (chap. II, § 2, théorème 5), le théorème que nous venons de démontrer n'impose aucune restriction à la nature topologique des espaces X_n. Cependant il est moins général que celui de Kolmogorov en ce sens qu'il relève d'une construction spéciale de mesures dans un produit d'espaces.

Revenons aux chaînes de Markov.

DÉFINITION. *On appelle chaîne de Markov d'espace des phases* $\{X, \mathfrak{B}\}$ *une famille de mesures* $\mathsf{P}^{(m)}(\cdot)$ *définies sur* $\{X \times X^\infty, \mathfrak{C}\}$ *dépendant d'une mesure arbitraire m sur* $\{X, \mathfrak{B}\}$ *comme d'un paramètre et dont les répartitions partielles sont définies par la formule*

$$\mathsf{P}^{(m)}\{\omega : x_k \in B_k, k = 0, \ldots, n\} = \\ = \int_{B_0} m(dx) \int_{B_1} \mathsf{P}_0(x, dy_1) \ldots \int_{B_{n-1}} \mathsf{P}_{n-1}(y_{n-1}, B_n), \tag{8}$$

où $\{\mathsf{P}_n(x, B), n = 0, 1, \ldots\}$ *est un système de noyaux stochastiques sur* $\{X, \mathfrak{B}\}$.

Les noyaux stochastiques $\mathsf{P}_n(x, B)$ s'appellent *probabilités de passage en une épreuve*, la mesure m *répartition initiale de la chaîne.* En fixant la mesure m on obtient une suite aléatoire à valeurs dans X appelée *processus markovien attaché à la répartition initiale m.*

On désignera par $\mathsf{P}^{(m)}_{t_1, t_2, \ldots, t_n}$ les répartitions finidimensionnelles de ce processus et par E_m l'espérance mathématique d'une fonction de processus par rapport à la mesure probabiliste $\mathsf{P}^{(m)}$.

Si m est concentrée en un point fixe x de l'espace des phases, ce point sera appelé *état initial du processus*, quant aux répartitions finidimensionnelles, à la mesure sur $\{X^\infty, \mathfrak{S}\}$ et à l'espérance mathématique d'une fonction de processus relativement à une mesure appropriée, on les désignera respectivement par $\mathsf{P}^{(x)}_{t_1, t_2, \ldots, t_n}$, $\mathsf{P}^{(x)}$ et E_x. Posons $(k < r)$

$$\mathsf{P}(k, x, r, B) = \int_X \mathsf{P}_k(x, dy_{k+1}) \int_X \mathsf{P}_{k+1}(y_{k+1}, dy_{k+2}) \ldots$$
$$\ldots \int_X \mathsf{P}_{r-2}(y_{r-2}, dy_{r-1}) \mathsf{P}_{r-1}(y_{r-1}, B).$$

Du point de vue analytique $\mathsf{P}(k, \cdot, r, \cdot)$ est un noyau stochastique, produit de convolution des probabilités de passage $\mathsf{P}_k * \mathsf{P}_{k+2} * \ldots$ $\ldots * \mathsf{P}_{r-1}$. On l'appelle aussi *probabilité de passage*. Plus exactement, $\mathsf{P}(k, x, r, B)$ est la probabilité de passage de l'état x à l'ensemble B dans l'intervalle de temps $]k, r[$. L'associativité du produit de convolution des noyaux entraîne

$$\mathsf{P}(k, x, s, B) = \int_X \mathsf{P}(k, x, r, dy) \mathsf{P}(r, y, s, B), \quad k < r < s, \tag{9}$$

c'est-à-dire l'équation de Kolmogorov-Chapman. La formule (7) donne

$$\mathsf{E}_m f(\xi(t_1), \xi(t_2), \ldots, \xi(t_s)) =$$
$$= \int m(dx) \int \mathsf{P}(0, x, t_1, dy_1) \int \mathsf{P}(t_1, y_1, t_2, dy_2) \times \ldots$$
$$\ldots \times \int f(y_1, y_2, \ldots, y_s) \mathsf{P}(t_{s-1}, y_{s-1}, t_s, dy_s). \tag{10}$$

Soit $\xi(m) = \xi(m, \omega)$ une fonction coordonnée sur $X \times X^\infty$:

$$\xi(m, \omega) = x_m, \; m = 0, 1, \ldots, \; \omega = (x_0, x_1, \ldots).$$

La formule (10) permet de préciser le sens probabiliste de la probabilité de passage. A cet effet calculons l'espérance mathématique conditionnelle de la fonction $f(\xi(s), \xi(s+1), \ldots, \xi(s+n))$ ($f(y_0, y_1, \ldots, y_n)$ est une fonction borélienne non négative de $n+1$ variables) par rapport à la tribu $\mathfrak{F}_{[0, t]}$ engendrée par les variables $\xi(0), \xi(1), \ldots, \xi(t)$, $t \leqslant s$. Notons Ψ l'espérance mathématique correspondante. Par définition Ψ est la seule variable aléatoire $\mathfrak{F}_{[0, t]}$-mesurable telle que pour toute fonction non négative

$g(x_0, x_1, \ldots, x_t)$ l'on ait

$$\mathsf{E}_m g(\xi(0), \xi(1), \ldots, \xi(t)) f(\xi(s), \xi(s+1), \ldots, \xi(s+n)) = \\ = \mathsf{E}_m g(\xi(0), \xi(1), \ldots, \xi(t)) \Psi.$$

Par ailleurs (10) entraîne

$$\mathsf{E}_m g(\xi(0), \xi(1), \ldots, \xi(t)) f(\xi(s), \xi(s+1), \ldots, \xi(s+n)) = \\ = \mathsf{E}_m g(\xi(0), \xi(1), \ldots, \xi(n)) \hat{f},$$

où

$$\hat{f} = \hat{f}(\xi(t)) = \int \mathsf{P}(t, \xi(t), s, dy_0) \int \mathsf{P}_s(y_0, dy_1) \times \cdots \\ \cdots \times \int f(y_0, y_1, \ldots, y_n) \mathsf{P}_{s+n-1}(y_{n-1}, dy_n).$$

Donc $\Psi = \hat{f}$.

La formule obtenue conduit aux conclusions suivantes:

THÉORÈME 4. *L'espérance mathématique conditionnelle de toute fonction non négative* $f(\xi(s), \xi(s+1), \ldots, \xi(s+n))$ *par rapport à* $\mathfrak{F}_{[0,t]}$ $(t \leqslant s)$ *ne dépend ni de la répartition initiale* m, *ni de la probabilité de passage antérieure à l'instant* t, *ni des valeurs* $\xi(0)$, $\xi(1), \ldots, \xi(t-1)$. *Elle est donnée par l'expression*

$$\mathsf{E}_m \{f(\xi(s), \xi(s+1), \ldots, \xi(s+n)) \mid \mathfrak{F}_{[0,t]}\} = \\ = \int \mathsf{P}(t, \xi(t), s, dy_0) \int \mathsf{P}_s(y_0, dy_1) \cdots \\ \cdots \int f(y_0, y_1, \ldots, y_n) \mathsf{P}_{s+n-1}(y_{n-1}, dy_n). \quad (11)$$

La répartition conditionnelle des variables $\xi(s)$, $\xi(s+1), \ldots$ $\ldots, \xi(s+n)$ dans $\{X^{n+1}, \mathfrak{B}_{n+1}\}$ par rapport à $\mathfrak{F}_{[0,t]}$ est confondue avec le produit direct des noyaux

$$\mathsf{P}(t, \xi(t), s, \cdot), \mathsf{P}_s(\cdot, \cdot), \ldots, \mathsf{P}_{s+n-1}(\cdot, \cdot).$$

En particulier, la probabilité de passage $\mathsf{P}(t, \xi(t), s, B)$ est confondue avec la probabilité conditionnelle que le système se trouve dans l'ensemble B à l'instant s sachant les états $\xi(0), \xi(1), \ldots, \xi(t)$. Cette probabilité dépend uniquement de l'état $\xi(t)$ au dernier instant connu et ne dépend ni des valeurs $\xi(0), \xi(1), \ldots, \xi(t-1)$, ni de m, ni des probabilités de passage $\mathsf{P}_1(\cdot, \cdot), \mathsf{P}_2(\cdot, \cdot), \ldots, \mathsf{P}_t(\cdot, \cdot)$. Cette propriété, qui, nous l'avons vu, traduit l'*absence de post-action*, est une caractéristique fondamentale de la chaîne markovienne.

REMARQUE. Soient donnés un espace probabilisable $\{X, \mathfrak{B}\}$ et un système de noyaux stochastiques $\mathsf{P}_n(x, B)$, $n = 0, 1, \ldots$, sur $\{X, \mathfrak{B}\}$. Il existe alors une chaîne de Markov pour laquelle $\mathsf{P}_n(x, B)$ sont des probabilités de passage en une épreuve. La démonstration de cette proposition et la construction de l'espace probabilisé correspondant sont données par le théorème 3.

Une chaîne de Markov est *homogène* si les probabilités de passage en une épreuve ne dépendent pas du temps :

$$\mathsf{P}_t(x, B) = \mathsf{P}(x, B).$$

Dans ce cas la probabilité de passage pendant l'intervalle séparé par les instants s et t ne dépend que de la longueur de cet intervalle : $\mathsf{P}(t, x, s, B) =$

$$= \int_X \mathsf{P}(x, dy_1) \int_X \mathsf{P}(y_1, dy_2) \ldots \int_X \mathsf{P}(y_{s-1}, B)\, \mathsf{P}(y_{s-2}, dy_{s-1}) = \mathsf{P}^{(s-t)}(x, B).$$

Dans le cas d'une chaîne homogène l'équation de Kolmogorov-Chapman se note

$$\mathsf{P}^{(s+m)}(x, B) = \int_X \mathsf{P}^{(s)}(x, dy)\, \mathsf{P}^{(m)}(y, B).$$

Supposons que la chaîne de Markov est homogène. La formule (10) nous dit que

$$\mathsf{E}_m f(\xi(s+1), \xi(s+2), \ldots, \xi(s+n)) = \mathsf{E}_{m_s} f(\xi(1), \xi(2), \ldots, \xi(n)), \quad (12)$$

où

$$m_s(B) = \int \mathsf{P}(0, x, s, B)\, m(dx) = \int \mathsf{P}^{(s)}(x, B)\, m(dx).$$

Si (12) ne dépend pas de s quelle que soit la fonction $f(\cdot)$, le processus markovien homogène attaché à la répartition initiale m donnée est par définition *stationnaire*. Pour qu'un processus soit stationnaire il faut et il suffit que la mesure m vérifie la condition

$$m(B) = \int \mathsf{P}^{(s)}(x, B)\, m(dx) \quad (13)$$

qui équivaut à la condition plus simple :

$$m(B) = \int \mathsf{P}(x, B)\, m(dx). \quad (14)$$

En effet, (14) est un cas particulier de (13). Si (14) est réalisée, alors

$$m(B) = \int \mathsf{P}(x, B) \int \mathsf{P}(y, dx)\, m(dy) = \int \mathsf{P}^{(2)}(y, B)\, m(dy) = \ldots \ldots = \int \mathsf{P}^{(s)}(y_s, B)\, m(dy_s).$$

Les mesures probabilistes m qui sont solutions de (14) sont dites *invariantes* ou, de façon plus détaillée, *mesures invariantes associées au noyau stochastique donné*.

Si donc au noyau stochastique donné est associée une mesure probabiliste invariante, il existe une répartition initiale d'une

chaîne de Markov à laquelle est associé un processus markovien stationnaire. La probabilité de passage en une épreuve est précisément le noyau donné.

Soient $\mathfrak{F}_t$ la plus petite tribu, par rapport à laquelle sont mesurables $\xi(0), \xi(1), \ldots, \xi(t)$ $(t = 0, 1, 2, \ldots)$, τ un instant markovien sur $\{\mathfrak{F}_t,\ t = 0, 1, \ldots\}$ prenant éventuellement la valeur ∞.

Posons le problème suivant. Soit $\xi(t)$ une chaîne de Markov homogène. Comment se comporte le processus $\xi_\tau(t) = \xi(t + \tau)$ pour $\tau < \infty$? Sous l'hypothèse $\xi(\tau) = x$, $\tau < \infty$, le processus aléatoire $\xi_\tau(t)$ se comportera semble-t-il exactement comme le processus markovien $\xi(t)$ sous l'hypothèse $\xi(0) = x$. Précisons et démontrons cette assertion qui exprime la *propriété d'être markovien fort* pour un processus. Soit $\Omega_\tau = \{\omega : \tau < \infty\}$. Posons $\mathsf{P}^{(\tau)}(x, A) = \mathsf{P}^{(x)}[\{\Omega_\tau \cap \cap (\xi(\tau) \in A)\}$. On a

$$\mathsf{P}^{\tau}(x, A) = \sum_{s=1}^{\infty} \mathsf{P}^{(x)}\{[\tau = s] \cap [\xi(s) \in A]\}.$$

D'où il suit que $\mathsf{P}^{(\tau)}(x, A)$ est une mesure sur $\mathfrak{B}$ et

$$\mathsf{P}^{\tau}(x, X) = \mathsf{P}^{(x)}\{\Omega_\tau\} \leqslant 1.$$

D'autre part, il existe un ensemble $B^{(s)} \in \mathfrak{B}^{(s)}$ tel que l'événement $\{\tau = s\}$ est équivalent à l'événement $\{\xi(0), \xi(1), \ldots, \xi(s)\} \in B^{(s)}$. Donc

$$\mathsf{P}^{(x)}\{[\tau = s] \cap [\xi(s) \in A]\} = \mathsf{P}^{(x)}\{(\xi(0), \xi(1), \ldots, \xi(s)) \in \in B^{(s)} \cap A^{(s)}\},$$

où $A^{(s)} = X \times X \times \ldots \times X \times A$ ($s - 1$ facteurs X) et de ce qu'il précède il suit que cette probabilité, de même que $\mathsf{P}^{(\tau)}(x, A)$, sont des fonctions $\mathfrak{B}$-mesurables.

Appelons $\mathfrak{F}_\tau$ la tribu induite par l'instant aléatoire τ.

Théorème 5. *Si $D \in \mathfrak{F}_\tau$ et $D \subset \Omega_\tau$, alors*

$$\mathsf{P}^{(x)}\{D \cap (\bigcap_{k=1}^{r} [\xi(t_k + \tau) \in A_k])\} = \int_X \mathsf{P}^{(y)}(\bigcap_{k=1}^{r} [\xi(t_k) \in A_k])\, \mathsf{P}^{(\tau)}(x, D, dy), \qquad (15)$$

où $\mathsf{P}^{(\tau)}(x, D, A) = \mathsf{P}^{(x)}(D \cap [\xi(\tau) \in A])$.

D é m o n s t r a t i o n. Vu que $D \subset \Omega_\tau$, on a

$$\mathsf{P}^{(x)}\{D \cap (\bigcap_{k=1}^{r} [\xi(t_k + \tau) \in A_k])\} = \sum_{s=1}^{\infty} \mathsf{P}^{(x)}\{D_s \cap (\sum_{k=1}^{r} [\xi(t_k + \tau) \in A_k])\},$$

où $D_s = D \cap [\tau = s]$. Soit $\chi(D_s)$ l'indicateur de l'événement D_s. En vertu des propriétés des probabilités conditionnelles pour une

chaîne de Markov (théorème 4), on a

$$\mathsf{P}^{(x)}\{D_s\cap(\bigcap_{k=1}^{r}[\xi(t_s+\tau)\in A_k])\}=$$

$$=\mathsf{E}_x\{\chi(D_s)\,\mathsf{P}^{(x)}(\bigcap_{k=1}^{r}[\xi(t_k+s)\in A_k]\mid\mathfrak{F}_s)\}=$$

$$=\mathsf{E}_x\{\chi(D_s)\,\mathsf{P}^{(\xi_s)}(\bigcap_{k=1}^{r}[\xi(t_k+s)\in A_k)\}.$$

La chaîne étant homogène, le second membre vaut

$$\int_{D_s}\mathsf{P}^{(y)}\{\bigcap_{k=1}^{r}[\xi(t_k)\in A_k]\}\,d\mathsf{P}^{(x)}=$$

$$=\int_X\mathsf{P}^{(y)}\{\bigcap_{k=1}^{r}[\xi(t_k)\in A_k]\}\,\mathsf{P}(s,\,x,\,D,\,dy),\qquad(16)$$

où $\mathsf{P}(s,\,x,\,D,\,\cdot)$ est une mesure définie sur $[X,\,\mathfrak{B}]$ par

$$\mathsf{P}(s,\,x,\,D,\,A)=\mathsf{P}^{(x)}\{D\cap[\tau=s]\cap[\xi(s)\in A]\}.$$

Si l'on considère encore la mesure

$$\mathsf{P}^{(\tau)}(x,\,D,\,A)=\mathsf{P}^{(x)}\{D\cap[\xi(\tau)\in A]\}=\sum_{s=1}^{\infty}\mathsf{P}(s,\,x,\,D,\,A)$$

et que dans (16) l'on étende la sommation à s, on obtient le théorème annoncé. ■

§ 6. Chaînes de Markov à nombre dénombrable d'états

Réductibilité et irréductibilité. Soit X un ensemble dénombrable ou fini. Par tribu d'ensembles mesurables de X on entendra toujours dans ce cas l'ensemble de toutes les parties de X. Toute fonction sur X sera alors mesurable.

On désignera les points de l'espace par les lettres $i,\,j,\,\ldots$ Soit une chaîne de Markov homogène à valeurs dans X. Elle est définie par des probabilités de passage en une épreuve $p(i,\,j)$, $i,\,j\in X$, dans des ensembles à un seul point j. La probabilité de passage en une épreuve dans un ensemble quelconque B s'exprime en fonction de $p(i,\,j)$ à l'aide de la formule évidente

$$\mathsf{P}(i,\,B)=\sum_{j\in B}p(i,\,j),$$

quant à l'intégration par rapport à la mesure associée au noyau stochastique $\mathsf{P}(i,\,B)$, elle se transforme en la sommation

$$\int_X f(j)\,\mathsf{P}(i,\,dj)=\sum_{j\in X}p(i,\,j)\,f(j).$$

La probabilité de passage dans un ensemble à un seul point j en n épreuves se note

$$\mathsf{P}^{(n)}(i, j) = \sum_{j_1, j_2, \ldots, j_{n-1} \in X} p(i, j_1)\, p(j_1, j_2) \cdots p(j_{n-1}, j). \tag{1}$$

Si l'on introduit une matrice $\mathsf{P}^{(n)}$ (à nombre fini ou infini de lignes) dont les éléments sont les probabilités de passage en n épreuves, $\mathsf{P}^{(n)} = \{\mathsf{P}^{(n)}(i, j)\}_{i, j \in X}$, la formule (1) entraîne que

$$\mathsf{P}^{(n)} = \mathsf{P}^n,$$

où P^n est la puissance n-ième de la matrice des probabilités de passage en une épreuve $\mathsf{P} = \mathsf{P}^{(1)}$. La matrice $\mathsf{P} = \{p(i, j)\}$ possède les propriétés suivantes:

$$\text{a) } p(i, j) \geqslant 0; \quad \text{b) } \sum_{j \in X} p(i, j) = 1. \tag{2}$$

Toute matrice P possédant les propriétés a) et b) est dite *matrice stochastique*. L'égalité $\mathsf{P}^{n+m} = \mathsf{P}^n \mathsf{P}^m$ implique

$$p^{(n+m)}(i, j) = \sum_{k \in X} p^{(n)}(i, k)\, p^{(m)}(k, j). \tag{3}$$

La formule (3) est une transcription de l'équation de Kolmogorov-Chapman ((9), § 5) dans le cas traité.

Définition. *Un état $j \in X$ est accessible à partir d'un état i si la probabilité de passage de i en j en un certain nombre d'épreuves est positive. Si j est accessible à partir de i et i à partir de j, les états i et j sont dits communicants. Par définition l'état i est toujours communicant avec lui-même.*

La communication de i et j sera notée $i \leftrightarrow j$. Si j est accessible à partir de i et k à partir de j, alors k l'est à partir de i. Ceci découle de l'inégalité

$$p^{(n+m)}(j, k) \geqslant p^{(n)}(i, j)\, p^{(m)}(j, k).$$

La relation $\leftrightarrow$ est une relation d'équivalence:

a) $i \leftrightarrow i$;

b) si $i \leftrightarrow j$, alors $j \leftrightarrow i$;

c) si $i \leftrightarrow j$ et $j \leftrightarrow k$, alors $i \leftrightarrow k$.

En effet, a) découle de ce que $p^{(0)}(i, i) = 1$; b) de la symétrie de i et j dans la définition des états communicants et enfin c) de

$$p^{(n+m)}(i, k) \geqslant p^{(n)}(i, j)\, p^{(m)}(j, k) > 0,$$
$$p^{(n_1+m_1)}(k, i) \geqslant p^{(n_1)}(k, j)\, p^{(m_1)}(j, i) > 0,$$

si $p^{(n)}(i, j) > 0$, $p^{(m_1)}(j, i) > 0$; $p^{(m)}(j, k) > 0$, $p^{(n_1)}(k, j) > 0$.

Une chaîne de Markov est susceptible d'être décomposée en classes disjointes X_α d'états communicants de la manière suivante. On prend un état quelconque i_1 et on note X_{i_1} l'ensemble de tous les états communicants avec i_1. La propriété c) de la relation $\leftrightarrow$ impli-

que que deux états quelconques de X_{i_1} sont communicants. Si X_{i_1} n'épuise pas X, on choisit un état $i_2 \bar{\in} X_{i_1}$ et on construit selon la même procédure la classe X_{i_2}. Les états i_1 et i_2 ne communiquant pas, les classes X_{i_1} et X_{i_2} n'auront pas d'éléments communs. On poursuit la construction des ensembles X_{i_k} jusqu'à épuisement de X. Les classes X_α possèdent les propriétés suivantes:

a) le nombre de classes X_α est au plus dénombrable;

b) tout élément n'appartient qu'à une et une seule classe X_α;

c) deux états quelconques de X_α sont communicants;

d) deux états quelconques de deux classes distinctes ne communiquent pas entre eux.

Les deux dernières propriétés s'énoncent encore: à partir d'un état donné i d'une classe X_α on peut en un certain nombre d'épreuves atteindre avec une probabilité positive un autre état de cette même classe. Il n'est pas exclu qu'un système se trouvant dans une classe la quitte, mais la probabilité qu'il y revienne est alors nulle.

DÉFINITION. *Une chaîne de Markov est irréductible si elle est constituée d'une seule classe d'états communicants. Un état i est essentiel (resp. non essentiel) si tout état j, accessible à partir de i, communique (resp. ne communique pas) avec i.*

On remarque sans peine qu'à partir d'un état essentiel on ne peut accéder qu'à des états essentiels. En effet, soient i essentiel et j accessible à partir de i. Si k est accessible à partir de j, il l'est à partir de i et i l'est à partir de k (puisque i est essentiel). Donc j est accessible à partir de k, c'est-à-dire j est essentiel.

D'où le corollaire: dans toute classe d'états communicants les états soit sont tous essentiels soit ne le sont pas.

Récurrence. Soit $\xi(n)$ l'état d'un système markovien à l'instant n. Désignons par $\tau_j = \tau_j(n)$ le nombre d'épreuves nécessaires au système markovien pour atteindre pour la première fois l'état j à partir de l'instant n. Donc $\tau_j(n)$ est défini par les relations

$$\xi(n+1) \neq j, \ \ldots, \ \xi(n+\tau_j-1) \neq j, \ \xi(n+\tau_j) = j.$$

Considérons la famille de tribus $\{\mathfrak{F}_{[n,t]}, \ t = 0, 1, \ldots\}$, où $\mathfrak{F}_{[n,t]}$ est la plus petite tribu par rapport à laquelle sont mesurables $\xi(n)$, $\xi(n+1), \ldots, \xi(n+t)$.

La quantité $\tau_j(n)$ est un instant markovien sur cette famille. Posons

$$f^{(s)}(i, j) = \mathsf{P}(\tau_j(n) = s \mid \xi(n) = i), \ s = 1, 2, \ldots,$$
$$f^{(0)}(i, j) = 0.$$

De plus

$$f^{(1)}(i, j) = p^{(1)}(i, j) = p(i, j).$$

L'homogénéité de la chaîne entraîne que les probabilités $f^{(s)}(i, j)$ ne dépendent pas de n. On les appelle *probabilités de premier accès*

à l'état j si $i \neq j$ et *probabilités de premier retour* à l'état i si $i = j$.

La somme

$$F(i, j) = \sum_{s=1}^{\infty} f^{(s)}(i, j) \qquad (i \neq j)$$

est la probabilité que le système quitte l'état i et accède jamais à l'état j. De façon analogue $F(i, i)$ est la probabilité que le système sorte de l'état i et y revienne au terme d'un nombre fini d'épreuves. Si $F(i, j) < 1$, alors τ_j est impropre.

DÉFINITION. *Un état i est dit récurrent si $F(i, i) = 1$ et transitoire si $F(i, i) < 1$.*

On établit sans peine une relation entre les probabilités de passage et les probabilités de premier accès, soit

$$p^{(n)}(i, j) = \sum_{s=1}^{n} f^{(s)}(i, j)\, p^{(n-s)}(j, j), \qquad n \geqslant 1. \tag{4}$$

On suppose que $p^{(0)}(i, j) = \delta_{ij} = 0$. En effet, soit τ_j l'instant de premier accès en j, compté à partir de l'instant initial. On a

$$\mathsf{P}^{(n)}(i, j) = \mathsf{P}^{(i)}\{\bigcup_{s=1}^{n} [\tau_j = s] \cap [\xi(n) = j]\} =$$

$$= \sum_{s=1}^{n} \mathsf{P}^{(i)}\{[\tau_j = s] \cap [\xi(n) = j]\} = \sum \mathsf{P}^{(i)}\{\tau_j = s\}\, \mathsf{P}^{(i)}\{\xi(n) = j\} =$$

$$= \sum_{s=1}^{n} f^{(s)}(i, j)\, \mathsf{P}^{(n-s)}(j, j),$$

d'où la formule (4). Signalons un cas particulier:

$$p^{(n)}(i, i) = \sum_{s=1}^{n} f^{(s)}(i, i)\, p^{(n-s)}(i, i), \tag{5}$$

qui s'écrit encore

$$f^{(n)}(i, i) = p^{(n)}(i, i) - \sum_{s=1}^{n-1} f^{(s)}(i, i)\, p^{(n-s)}(i, i).$$

La dernière expression permet de calculer de proche en proche la probabilité de retour si sont connues les probabilités de passage. On remarquera que pour calculer les probabilités de retour à l'état i il suffit seulement de connaître les probabilités de passage en cet état.

Soient $P_{ij}(z)$ et $F_{ij}(z)$ fonctions génératrices des suites $\{p^{(n)}(i, j),\ n = 0, 1, 2, \ldots\}$, $\{f^{(n)}(i, j),\ n = 0, 1, 2, \ldots\}$:

$$P_{ij}(z) = \sum_{n=0}^{\infty} p^{(n)}(i, j)\, z^n, \quad F_{ij}(z) = \sum_{n=0}^{\infty} f^{(n)}(i, j)\, z^n.$$

La formule (5) donne

$$P_{ii}(z) = p^{(0)}(i,\ i) + \sum_{n=1}^{\infty} \sum_{k=1}^{n} f^{(k)}(i,\ i)\, z^k p^{(n-k)}(i,\ i)\, z^{n-k} =$$

$$= 1 + \sum_{k=1}^{\infty} \sum_{n=k}^{\infty} f^{(k)}(i,\ i)\, z^k p^{(n-k)}(i,\ i)\, z^{n-k} = 1 + \sum_{k=1}^{\infty} f^{(k)}(i,\ i)\, z^k P_{ii}(z)$$

ou

$$P_{ii}(z) = 1 + P_{ii}(z)\, F_{ii}(z).$$

Le changement de l'ordre de sommation est licite, car les séries étudiées sont absolument convergentes lorsque $|z| \leqslant 1$. La dernière formule peut encore s'écrire

$$P_{ii}(z) = \frac{1}{1 - F_{ii}(z)}. \tag{6}$$

De façon analogue on déduit de (4)

$$P_{ij}(z) = P_{ij}(z)\, F_{ij}(z),\ i \neq j. \tag{7}$$

Soit z un réel et $z \uparrow 1$. Les fonctions $P_{ii}(z)$ et $F_{ii}(z)$ sont monotones croissantes et de plus, en vertu du théorème d'Abel, $\lim\limits_{z \uparrow 1} F_{ii}(z) = F_{ii}(1) = F(i,\ i)$. Posons

$$\lim_{z \uparrow 1} P_{ii}(z) = G(i,\ i) = P_{ii}(1).$$

La relation (6) entraîne le

THÉORÈME 1. *Un état i est récurrent si* $G(i,\ i) = \sum\limits_{n=0}^{\infty} p^{(n)}(i, i) = \infty$ *et transitoire si* $G(i,\ i) = \sum\limits_{n=0}^{\infty} p^{(n)}(i,\ i) < \infty$. *Dans le cas transitoire*

$$G(i,\ i) = \frac{1}{1 - F(i,\ i)}.$$

THÉORÈME 2. *Si des états i et j sont communicants, ils sont simultanément récurrents ou transitoires.*

D é m o n s t r a t i o n. Comme $i \leftrightarrow j$, il existe un m_1 et un m_2 tels que

$$p^{(m_1)}(i,\ j) > 0, \quad p^{(m_2)}(j,\ i) > 0.$$

D'autre part,

$$p^{(m_1+m_2+n)}(j,\ j) \geqslant p^{(m_2)}(j,\ i)\, p^{(n)}(i,\ i)\, p^{(m_1)}(i,\ j),$$

donc

$$\sum_{n=m_1+m_2}^{\infty} p^{(n)}(j,\ j) \geqslant p^{(m_2)}(j,\ i)\, p^{(m_1)}(i,\ j) \sum_{n=0}^{\infty} p^{(n)}(i,\ i)$$

et la série $G(j,\ j)$ est divergente si l'est $G(i,\ i)$. En permutant i et j on conclut que $G(i,\ i)$ et $G(j,\ j)$ sont simultanément finis ou infinis. ■

Donc pour une chaîne de Markov la propriété de récurrence est moins une propriété d'état qu'une caractéristique d'états communicants.

L'intuition nous suggère que les retours au bout d'un intervalle de temps infini dans un état récurrent ont lieu une infinité de fois et dans un état transitoire seulement un nombre fini de fois. La démonstration ne soulève aucune difficulté.

Soient $Q_j(m)$ l'événement : {le système accède à l'état j au moins m fois}, τ_j le nombre d'épreuves précédant la première arrivée à l'état j. Alors

$$Q_j(m) = \bigcup_{n=1}^{\infty} Q_j(m) \cap \{\tau_j = n\}.$$

Soit $q_{ij}(m)$ la probabilité de l'événement $Q_j(m)$ si $\xi(0) = i$. On a

$$q_{ij}(m) = \sum_{n=1}^{\infty} \mathsf{P}(Q_j(m) \cap [\tau_j = n] \mid \xi(0) = i) =$$
$$= \sum_{n=1}^{\infty} \mathsf{P}^{(i)}(Q_j(m) \mid \tau_j = n)\, \mathsf{P}^{(i)}(\tau_j = n \mid \xi(0) = i) =$$
$$= \sum_{n=1}^{\infty} f^{(n)}(i, j)\, \mathsf{P}^{(i)}(Q_j(m) \mid \tau_j = n).$$

On vérifie immédiatement que

$$\mathsf{P}^{(i)}(Q_j(m) \mid \tau_j = n) = \mathsf{P}^{(j)}(Q_j(m-1)) = q_{jj}(m-1).$$

Donc

$$q_{ij}(m) = F(i, j)\, q_{jj}(m-1). \tag{8}$$

Soit $q_{ij} = q_{ij}(\infty)$ la probabilité que sorti de l'état i le système entre une infinité de fois dans l'état j. Comme $q_{ij} = \lim_{m\to\infty} q_{ij}(m)$, on déduit de (8) que

$$q_{ij} = F(i, j)\, q_{jj}. \tag{9}$$

THÉORÈME 3. *Si j est un état récurrent, alors $q_{ij} = F(i, j)$ et en particulier $q_{jj} = 1$; si j est transitoire, alors $q_{ij} = 0$ pour tout i.*

D é m o n s t r a t i o n. Si $F(j, j) < 1$, en faisant $i = j$ dans (9), on obtient $q_{jj} = 0$, et de la même égalité on déduit que $q_{ij} = 0$. Si $F(j, j) = 1$, alors de (8) il vient que $q_{jj}(m) = [F(j, j)]^{m-1} = 1$, d'où $q_{jj} = 1$. De (9) il résulte alors : $q_{ij} = F(i, j)$. ■

Soit $F(i, j) = 1$. La propriété d'être markovien fort (cf. § 5, théorème 5) implique

$$\mathsf{P}^{(i)}(B \cap (\bigcap_{k=1}^{r} \{\xi(\tau_j + t_k) = j_k\})) = \mathsf{P}^{(i)}(B)\, \mathsf{P}^{(j)}(\bigcap_{k=1}^{r} \{\xi(t_k) = j_k\})$$

pour tout $B \in \mathfrak{F}_{\tau_j}$. Cette relation entraîne le

THÉORÈME 4. *Si* $F(i, j) = 1$, *le processus aléatoire* $\xi'(t) = \xi(\tau_j + t)$ ($\xi(0) = i$) *est stochastiquement équivalent au processus* $\xi(t)$ *à état initial* $\xi(0) = j$ *et ne dépend pas de la tribu* $\mathfrak{F}_{\tau_j}$.

COROLLAIRE *Soient* $\xi(0) = i$, i *est un état récurrent*, ξ_1 *le nombre d'épreuves antérieures au premier retour en* i, ξ_2 *le nombre d'épreuves entre le premier et le second retour en* i, *etc.*

Les variables aléatoires $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *sont équi-réparties et indépendantes.*

THÉORÈME 5. *Si l'état* i *est récurrent et* $F(i, j) > 0$, *en quittant* i *le système passe à l'état* j *une infinité de fois* ($q_{ij} = 1$), *et* $F(j, i) > 0$. *En particulier*, $F(i, j) = 1$.

Le théorème 3 dit que le nombre de retours à l'état i est infini. Soit C_k l'événement : {le système passe par l'état j entre le $(k-1)$-ième et le k-ième séjour dans l'état i}. Le processus étant markovien fort, les événements C_k sont mutuellement indépendants et admettent la même probabilité. Comme $\bigcup_{k=1}^{\infty} C_k$ est la probabilité que le système passe jamais à l'état j, $\mathsf{P}(C_k) > 0$ et $\sum_{k=1}^{\infty} \mathsf{P}(C_k) = \infty$. Le théorème de Borel-Cantelli dit que C_k se réalise presque sûrement une infinité de fois. Bien plus, après avoir atteint l'état j le système passera une infinité de fois à l'état i. ■

COROLLAIRE 1. *A partir d'un état récurrent seuls sont accessibles des états récurrents. Les états récurrents sont essentiels.*

Ce corollaire précise le théorème 2 obtenu précédemment à l'aide des fonctions génératrices.

COROLLAIRE 2. *Dans une classe d'états communicants contenant un état récurrent, tous les autres états sont récurrents et tout système évoluant dans cette classe atteint presque sûrement tôt ou tard tous les autres états de la classe et ce une infinité de fois.*

Les classes d'états communicants récurrents sont appelées *classes récurrentes.*

Posons

$$G(i, j) = \sum_{n=0}^{\infty} p^{(n)}(i, j).$$

Cette série a été interprétée pour le cas $i = j$. Etablissons la relation suivante dans le cas générale :

$$\lim_{N\to\infty} \frac{\sum_{n=1}^{N} p^{(n)}(i, j)}{\sum_{n=0}^{N} p^{(n)}(j, j)} = F(i, j). \tag{10}$$

La démonstration est basée sur la formule (4). En faisant $n = 1, 2, \ldots, N$ dans (4) et en sommant les égalités obtenues on aura

$$\sum_{n=1}^{N} p^{(n)}(i, j) = \sum_{n=1}^{N} \sum_{s=0}^{n-1} f^{(n-s)}(i, j)\, p^{(s)}(j, j) =$$
$$= \sum_{s=0}^{N-1} \sum_{n=s+1}^{N} f^{(n-s)}(i, j)\, p^{(s)}(j, j) = \sum_{s=0}^{N-1} p^{(s)}(j, j)\, F_{N-s},$$

où $F_{N-s} = \sum_{n=1}^{N-s} f^{(n)}(i, j)$ et $F_N \to F(i, j)$ lorsque $N \to \infty$. Donc

$$\frac{\sum_{n=1}^{N} p^{(n)}(i, j)}{\sum_{n=0}^{N} p^{(n)}(j, j)} = \sum_{s=0}^{N} F_{N-s} \frac{p^{(s)}(j, j)}{\sum_{n=0}^{N} p^{(n)}(j, j)}.$$

La véracité de la formule (10) découle maintenant du

LEMME 1. *Si* b_n, $n = 0, 1, \ldots$, *est une suite de nombres non négatifs et* $\frac{b_N}{\sum_{s=0}^{N} b_s} \to 0$, *alors pour toute suite convergente* c_n, $n = 1, 2, \ldots$, *on a*

$$\lim_{N \to \infty} \frac{\sum_{k=0}^{N} b_k c_{N-k}}{\sum_{k=0}^{N} b_k} = \lim_{n \to \infty} c_n.$$

Démonstration. Si $c = \lim c_n$, alors

$$\frac{\sum_{k=0}^{N} b_k c_{N-k}}{\sum_{k=0}^{N} b_k} - c =$$
$$= \frac{\sum_{k=0}^{N-n} b_k (c_{N-k} - c)}{\sum_{k=0}^{N} b_k} - c \frac{\sum_{k=N-n+1}^{N} b_k}{\sum_{k=0}^{N} b_k} + \frac{\sum_{k=N-n+1}^{N} b_k c_{N-k}}{\sum_{k=0}^{N} b_k}. \quad (11)$$

Si l'indice n est choisi tel que pour $n' \geqslant n$ on ait $|c - c_{n'}| < \varepsilon$, $\forall \varepsilon > 0$, le premier terme du second membre de (11) est inférieur à ε. Pour n fixe le second et le troisième terme tendent également vers 0 lorsque $N \to \infty$, car les c_n sont bornés. D'où le lemme et l'égalité

(10) puisque les conditions du lemme sont valables pour le cas étudié, car les $p^{(n)}(j, i)$ sont bornés. ■

La formule (10) implique le

THÉORÈME 6. *$G(i, j) = \infty$ à l'intérieur d'une classe récurrente; si j est transitoire, $G(i, j) < \infty$ pour tous les i.*

En effet, si j est un état transitoire, le dénominateur du premier membre de (10) tend vers une limite finie, donc le numérateur tendra aussi vers une limite finie. Si au contraire j est récurrent, le dénominateur tend vers l'infini. Si $F(i, j) > 0$, le numérateur tend également vers l'infini.

Périodicité. On remarquera que si $p^{(n)}(i, i) > 0$, alors $p^{(kn)}(i, i) > 0$. En effet, $p^{(kn)}(i, i) \geqslant p^{(n)}(i, i)\, p^{(n)}(i, i) \ldots \ldots p^{(n)}(i, i)$. Soit $d(i)$ le P.G.C.D. de tous les n tels que $p^{(n)}(i, i) > > 0$. Si $p^{(n)}(i, i) = 0$ pour tous les $n \geqslant 1$, on admettra alors que $d(i) = \infty$.

THÉORÈME 7. *Si $i \leftrightarrow j$, alors $d(i) = d(j)$*

D é m o n s t r a t i o n. A noter que si $i \leftrightarrow j$, $d(i)$ et $d(j)$ sont finis. Soit $p^{(s)}(i, i) > 0$. Il existe un $n > 0$ et un $m > 0$ tels que $p^{(n)}(i, j) > 0$ et $p^{(m)}(j, i) > 0$, de sorte que $p^{(n+m+s)}(j, j) \geqslant \geqslant p^{(m)}(j, i)\, p^{(s)}(i, i)\, p^{(n)}(i, j) > 0$. De façon analogue $p^{(n+m+ks)}(j, j) > 0$. Donc $d(j)$ divise $(n + m + 2s) - - (n + m + s) = s$. D'où il suit que $d(j) \leqslant d(i)$. Mutatis mutandis on obtient $d(i) = d(j)$. ■

COROLLAIRE. *Dans toute classe d'états communicants $d(i)$ est constant.*

En particulier, $d = d(i)$ ne dépend pas de l'état dans une chaîne de Markov irréductible.

DÉFINITION. *Une chaîne de Markov irréductible est apériodique si $d = 1$, périodique et de période d si $d > 1$.*

THÉORÈME 8. *Si $d(i) < \infty$, il existe un n_0 tel que pour $n > n_0$*

$$p^{(nd(i))}(i, i) > 0.$$

D é m o n s t r a t i o n. Soit n_k $(k = 1, 2, \ldots, s)$ une suite de nombres tels que $p^{(n_k)}(i, i) > 0$ et que le P.G.C.D. des nombres $n_1, n_2, \ldots, n_s$ soit égal à $d(i)$. En vertu du lemme 4, § 4, il existe un n_0 tel que pour $n \geqslant n_0$ on ait $nd(i) = \sum_{k=1}^{s} c_k n_k$. Donc

$$p^{(nd(i))}(i, i) \geqslant [p^{(n_1)}(i, i)]^{c_1} \ldots [p^{(n_s)}(i, i)]^{c_s} > 0. \; ■$$

COROLLAIRE. *Si $p^{(m)}(j, i) > 0$, pour tous les n assez grands*

$$p^{(m+nd(i))}(j, i) > 0.$$

En effet,

$$p^{(m+nd(i))}(j, i) \geqslant p^{(m)}(j, i)\, p^{(nd(i))}(i, i).$$

On a souvent intérêt en étudiant les chaînes de Markov à considérer d'abord des chaînes apériodiques et ensuite à généraliser les résultats obtenus aux chaînes périodiques.

Montrons que la période d'un état se calcule à l'aide de la probabilité du premier retour.

LEMME 2. *La période de l'état i est confondue avec le P.G.C.D. de tous les n tels que* $f^{(n)}(i, i) > 0$.

Démonstration. Soient Z_N et Z'_N les ensembles des $n \leqslant N$ tels que $p^{(n)}(i, i) > 0$ et $f^{(n)}(i, i) > 0$ respectivement, et d_N et d'_N leurs P.G.C.D. De toute évidence $Z'_N \subset Z_N$, donc $d'_N \geqslant d_N$. En outre, $d'_1 = d_1$. Supposons qu'existe un N tel que $d'_n = d_n$ pour $n \leqslant N$ et $d'_{N+1} > d_{N+1}$. Alors $f^{(N+1)}(i, i) = 0$ et $p^{N+1}(i, i) > 0$, L'égalité

$$p^{(N+1)}(i, i) = f^{(N+1)}(i, i) + \sum_{k=1}^{N} f^{(k)}(i, i)\, p^{(N+1-k)}(i, i)$$

entraîne pour un s, $0 < s \leqslant N$,

$$f^{(s)}(i, i)\, p^{(N+1-s)}(i, i) > 0,$$

c'est-à-dire s et $N+1-s$ sont divisibles par d_N et par suite $N+1$ est divisible par d_N, ce qui contredit la relation $d_{N+1} < d'_{N+1} = d_N$. ■

THÉORÈME 9. *Toute classe K d'états communicants de période* $d < \infty$ *est susceptible d'être partitionnée en d sous-ensembles* $K_0, K_1, \ldots, K_{d-1}$ *deux à deux disjoints tels que, en une épreuve, de* K_s $(s < d-1)$ *on ne puisse passer qu'en* K_{s+1} *et de* K_{d-1} *en* K_0. *De plus, si* $i \in K_r$, $j \in K_s$, *il existe* $N = N(i, j)$ *tel que* $p^{(nd+s-r)}(i, j) > 0$ *pour* $n > N$.

Démonstration. Soit K_0 l'ensemble de tous les états j tels que pour au moins un k entier positif on ait $p^{(kd)}(i, j) > 0$, où i est un état choisi arbitrairement dans K. Alors $i \in K_0$. Comme i et j communiquent, il existe m tel que $p^{(m)}(j, i) > 0$. Le nombre m est multiple de d. En effet, $p^{(kd+m)}(i, i) \geqslant p^{(kd)}(i, j)\, p^{(m)}(j, i) > 0$, donc $kd+m$ est divisible par d. Comme m est divisible par d, si, au lieu de l'état i, on considère dans la définition de K_0 un j quelconque tel que $p^{(kd)}(i, j) > 0$ pour k quelconque, alors K_0 ne change pas. Soient maintenant K_1 l'ensemble des $j \in K$ tels que $\sum_{i \in K_0} p(i, j) > 0$, K_2 l'ensemble des j tels que $\sum_{i \in K_1} p(i, j) > 0$, $j \in K$, etc. Par définition de l'ensemble K_s on a $K_{rd+s} \subset K_s$ quels que soient r et s. D'autre part, si $j \in K_s$, il existe $j_0, j_1, \ldots, j_s = j$ tels que $j_r \in K_r$, $r \leqslant s$, et $p(j_{r-1}, j_r) > 0$, c'est-à-dire $p^{(s-r)}(j_r, j) > 0$. La réciproque est

vraie: si $p^{(s-r)}(j_r, j) > 0$, $j_r \in K_r$, $j \in K$, alors $j \in K_s$ ($j_r \to j_{r+1}$, $j_{r+1} \to j$ et $j_r \leftrightarrow j \Rightarrow j_{r+1} \leftrightarrow j_r$). Reste à s'assurer que les classes K_r et K_s sont disjointes, $0 \leqslant r < s < d$. Soit en effet $j \in K_r$, $j \in K_s$. Il existerait alors i_1 et $i_2 \in K_0$ tels que $p^{(r)}(i_1, j) > 0$ et $p^{(s)}(i_2, j) > 0$. Comme i_2 et j sont communicants, $p^{(m)}(j, i_2) > 0$ pour un m. Donc $p^{(m+s)}(i_2, i_2) \geqslant p^{(s)}(j, j)\, p^{(m)}(j, i_2) > 0$, et $m + s$ est divisible par d, c'est-à-dire $m = kd - s$, où s est un entier. On aurait alors

$$0 < p^{(r)}(i_1, j)\, p^{(m)}(j, i_2) \leqslant p^{(kd-s+r)}(i_1, i_2),$$

or, on l'a vu plus haut, ceci est impossible puisqu'on ne peut passer de i_1 en i_2 (i_1, $i_2 \in K_0$) qu'en un nombre d'épreuves multiple de d. Soit d'autre part $i \in K_r$ et $j \in K_s$. Il existe m tel que $p^{(m)}(i, j) > 0$. Alors $m = k_0 d + (s - r)$. Par ailleurs, en vertu du théorème 8 on a $p^{(nd)}(i, i) > 0$ pour tous les $n \geqslant n_0(i)$. Donc $p^{((n+k_0)d+s-r)}(i, j) \geqslant p^{(nd)}(i, i)\, p^{(k_0 d+s-r)}(i, j) > 0$ pour tous les $n > n_0(i)$. ■

Appelons les ensembles $K_0, K_1, \ldots, K_{d-1}$ *sous-classes de la classe périodique des états communicants.*

THÉORÈME 10. *Soient $p^{(n)}(i, j)$ les probabilités de passage d'une chaîne de Markov récurrente irréductible apériodique. Désignons par m_i le nombre moyen d'épreuves avant le premier retour à l'état i*

$$m_i = \sum_{n=1}^{\infty} n f^{(n)}(i, i).$$

Alors pour tout j

$$\lim_{n\to\infty} p^{(n)}(j, i) = \frac{1}{m_i}. \tag{12}$$

D é m o n s t r a t i o n. Soit τ_0 le nombre d'épreuves avant le premier retour à l'état i, τ_1 le nombre d'épreuves entre le premier et le second retour dans cet état, etc. Le corollaire du théorème 4 nous dit que $\tau_0, \tau_1, \ldots, \tau_n, \ldots$ sont mutuellement indépendants, équirépartis, prennent des valeurs entières non inférieures à l'unité, et de plus

$$\mathsf{P}\{\tau_k = n\} = f^{(n)}(i, i), \quad \sum_{n=1}^{\infty} f^{(n)}(i, i) = 1, \quad \mathsf{E}\tau_k = m_i.$$

Soit un processus de renouvellement dans lequel τ_n est la durée du n-ième renouvellement. Le rôle de p_n et de $G(n)$ incombe maintenant à $f^{(n)}(i, i)$ et $p^{(n)}(i, i)$ respectivement. La chaîne étant apériodique, le renouvellement le sera également en vertu du théorème 2.

Le théorème 2, § 4, entraîne

$$\lim_{n\to\infty} p^{(n)}(i, i) = \frac{1}{m_i};$$

on reconnaît ici un cas particulier de (12) pour $j = i$. La généralisation ne soulève aucune difficulté. La formule (4) donne

$$\frac{p^{(n)}(j, i)}{\sum_{k=1}^{n} f^{(k)}(j, i)} = \sum_{k=1}^{n} \widetilde{f}^{(k)}(j, i)\, p^{(n-k)}(i, i), \widetilde{f}^{(n)}(j, i) = \frac{f^{(n)}(j, i)}{\sum_{k=1}^{n} f^{(k)}(j, i)}.$$

En remarquant que

$$f^{(n)}(j, i) \to 0, \quad \sum_{k=1}^{n} f^{(k)}(j, i) \to 1 \text{ lorsque } n \to \infty$$

(en vertu de la récurrence et de l'irréductibilité de la chaîne) et en appliquant le lemme 1, on obtient la généralisation de (12). ■

Le théorème démontré s'appelle souvent *théorème ergodique des chaînes de Markov.*

Théorème 11. *Si une chaîne de Markov récurrente irréductible est périodique de période d, alors*

$$\lim_{n\to\infty} p^{(nd)}(i, i) = \frac{d}{m_i}. \tag{13}$$

Si K_s sont les sous-classes définies dans le théorème 9 *et $i \in K_r$, $j \in K_s$, alors*

$$\lim_{n\to\infty} p^{(nd+l)}(i, j) = \begin{cases} \frac{d}{m_j}, & l = s - r \pmod{d}, \\ 0, & l \neq s - r \pmod{d}. \end{cases} \tag{14}$$

Démonstration. Le lemme 2 implique que la période d'une chaîne irréductible de Markov est confondue avec la période du processus de renouvellement défini dans la démonstration du théorème précédent. Donc (13) découle directement du corollaire du théorème 2, § 4. Le théorème 9 donne: $p^{(nd+l)}(i, j) = 0$ pour $i \in K_r$, $j \in K_s$ et $l \neq s - r \pmod{d}$. Donc si $r < s$ par exemple, alors

$$p^{(nd+s-r)}(i, j) = \sum_{k=0}^{n} f^{(kd+s-r)}(i, j)\, p^{(n-k)d}(j, j).$$

On achève de démontrer la formule (14) exactement comme le théorème 10, en se référant au lemme 1. ■

Définition. *Un état récurrent j est*

$$\textit{nul si} \lim p^{(nd_j)}(j, j) = 0,$$

$$\textit{positif si} \lim p^{(nd_j)}(j, j) > 0.$$

Dans une classe récurrente les états sont tous simultanément positifs ou nuls. En effet, si $i \leftrightarrow j$, l'inégalité

$$p^{(m+nd_j+s)}(i, i) \geqslant p^{(m)}(i, j)\, p^{(nd_j)}(j, j)\, p^{(s)}(j, i),$$

où m et s sont tels que $p^{(m)}(i, j) > 0$, $p^{(s)}(j, i) > 0$, entraîne que

$$\lim p^{(nd)}(i, i) \geqslant \lim p^{(nd)}(j, j), \quad d = d_i = d_j.$$

Mutatis mutandis on démontre la proposition annoncée.

Résumons ce qui précède.

THÉORÈME 12.

a) *Une condition nécessaire et suffisante pour qu'un état j soit transitoire est que* $G_{jj} = \sum_{n=1}^{\infty} p^{(n)}(j, j) < \infty$. *Cela étant,*

$$G_{ij} = \sum_{n=1}^{\infty} p^{(n)}(i, j) \leqslant G_{jj} < \infty, \quad \lim_{n \to \infty} p^{(n)}(i, j) = 0, \quad \forall i,$$

b) *Soit j un état récurrent de période d et de durée moyenne de retour m_j. Si i est accessible à partir de j, alors i est également récurrent de même période d, nul ou positif en même temps que j, et il existe un k, $0 \leqslant k < d$, dépendant uniquement de i et j et tel que*

$$\lim p^{(md+r)}(i, j) = \begin{cases} \dfrac{d}{m_j} & \text{pour } r = k, \\ 0 & \text{pour } r \not\equiv k \pmod{d}, \end{cases} \tag{15}$$

c) *Si i et j appartiennent à une même classe récurrente,*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} p^{(n)}(i, j) = \frac{1}{m_j}. \tag{16}$$

La dernière proposition est une conséquence de b). Par ailleurs, contrairement à la proposition b) la formule (16) ne fait pas de distinction entre les classes d'états apériodiques et périodiques.

Appelons *positive* (resp. *nulle*) une chaîne de Markov récurrente irréductible à états positifs (resp. nuls).

Critères de récurrence. Répartitions stationnaires. La récurrence (nulle ou positive) de la chaîne de Markov est étroitement liée aux solutions non triviales du système linéaire homogène

$$\sum_{j \in I} p(i, j) x_j = x_i, \quad i \in I, \tag{17}$$

et de son transposé

$$\sum_{j \in I} p(j, i) x_j = x_i, \quad i \in I. \tag{18}$$

Si le système (17) possède une solution non négative et sommable, c'est-à-dire $x_i \geqslant 0$, $\sum x_i < \infty$, on admet que $\sum x_i = 1$ et une telle solution peut être traitée comme une répartition initiale invariante $x_i = \mathsf{P}\{\xi(0) = i\} = \mathsf{P}\{\xi(1) = i\} = \ldots$ engendrée par un processus markovien stationnaire. D'autre part, l'existence d'un pro-

cessus markovien stationnaire dont les probabilités de passage sont données équivaut à la donnée d'une solution sommable non négative du système (17).

S'agissant du système transposé (18), l'existence d'une solution non triviale $x_i = c$ coule de source. Pour une chaîne de Markov récurrente il est caractéristique que (18) n'admet pas d'autres solutions non triviales non négatives. Bien plus, on a le

THÉORÈME 13. *Une chaîne de Markov irréductible est récurrente si et seulement si le système d'inéquations*

$$\sum_{j\in I} p(i,\ j)\, x_j \leqslant x_i,\quad i\in I, \tag{19}$$

ne possède pas de solutions non négatives distinctes des solutions de la forme $x_i = c,\ i \in I$.

Démonstration. Soient une chaîne récurrente, $x_i \geqslant 0$ et x_i $(i \in I)$ solutions du système (19). Prenons $x_l > 0$ (le cas échéant, tous les $x_i \equiv 0$). De (19) il suit

$$x_i \geqslant \sum_{j\in I} p(i,\ j) \sum_{k\in I} p(j,\ k)\, x_k = \sum_{k\in I} p^{(2)}(i,\ k)\, x_k,$$

et par récurrence

$$x_i \geqslant \sum_{k\in I} p^{(n)}(i,\ k)\, x_k.$$

Pour tout i il existe n tel que $p^{(n)}(i,\ l) > 0$, donc

$$x_i \geqslant p^{(n)}(i,\ l)\, x_l > 0.$$

Par suite

$$x_i > 0,\ \forall i \in I.$$

Posons $y_i = \dfrac{x_i}{x_l}$, où l est un état choisi arbitrairement. On a

$$y_l \geqslant \sum_{j\in I}^{l} p(i,\ j)\, y_j \geqslant p(i,\ l) + \sum_{j\neq l} p(i,\ j)\, y_j.$$

En appliquant cette inégalité aux y_j du deuxième membre on obtient

$$y_l \geqslant p(i,\ l) + \sum_{j\neq l} p(i,\ j)\, p(j,\ l) + \sum_{j\neq l}\sum_{k\neq l} p(i,\ j)\, p(j,\ k)\, y_k =$$

$$= f^{(1)}(i,\ l) + f^{(2)}(i,\ l) + \sum_{k\neq l} {}_l p^{(2)}(i,\ k)\, y_k,$$

où ${}_l p^{(2)}(i, k) = \sum_{j\neq l} p(i,\ j)\, p(j,\ k)$ est la probabilité, en quittant l'état i, d'accéder en deux épreuves à l'état k et ce sans passer par l'état l. En répétant la procédure précédente on est conduit à l'inégalité

$$y_i \geqslant \sum_{n=1}^{N} f^{(n)}(i,\ l) + \sum_{k\neq l} {}_l p^{(N)}(i,\ k)\, y_k,$$

où ${}_l p^{(N)}(i, k)$ est la probabilité d'aller en N épreuves de l'état i à l'état k et ce sans passer par l'état l. En supposant $N \to \infty$, on obtient

$$y_i \geqslant \sum_{n=1}^{\infty} f^{(n)}(i, l) = 1$$

c'est-à-dire $x_i \geqslant x_l$.

i et l étant arbitraires, on trouve : $x_i = x_l =$ const, c'est-à-dire le système d'inéquations (19) ne possède pas de solutions non négatives distinctes de $x_i = c$, $i \in I$.

Supposons que la chaîne comporte au moins un état non récurrent (l'irréductibilité de la chaîne n'est pas utilisée). Posons $x_l = 1$, $x_i = F(i, l)$ pour $i \neq l$, où l est un état récurrent quelconque. On remarquera que $F(i, l) = 1$ pas pour tous les i, $i \neq l$. En effet, dans le cas contraire on aurait eu

$$F(l, l) = \sum_{k \neq l} p(l, k) F(k, l) + p(l, l) = \sum_{k \in I} p(l, k) = 1,$$

ce qui contredit la non-récurrence de l'état l. Donc les nombres non négatifs x_i définis plus haut ne sont pas tous égaux entre eux. Pour $i \neq l$ on a

$$x_i = F(i, l) = \sum_{k \neq l} p(i, k) F(k, l) + p(i, l) = \sum_{k \in I} p(l, k) x_k$$

et

$$x_l = 1 > F(l, l) = \sum_{k \in I} p(l, k) x_k,$$

c'est-à-dire $\{x_i, i \in I\}$ est solution non négative et non constante du système (19). ■

Penchons-nous maintenant sur le lien entre l'existence de répartitions initiales invariantes et les propriétés de récurrence d'une chaîne de Markov, c'est-à-dire sur la résolubilité du système (17) pour une chaîne récurrente.

THÉORÈME 14. *Soit une chaîne de Markov irréductible et récurrente. Le système d'équations* (17) *admet au plus une solution vérifiant les conditions*

$$\sum_{i \in I} |x_i| < \infty, \quad \sum_{i \in I} x_i = 1. \tag{20}$$

Si une chaîne est récurrente et positive, la solution du système (17) *vérifiant* (20) *est de la forme*

$$x_i = v_i = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} p^{(n)}(j, i). \tag{21}$$

Si la chaîne est récurrente et nulle, l'unique solution (absolument sommable) du système (17) *est triviale* ($x_i = 0$).

D é m o n s t r a t i o n. Prouvons tout d'abord que le système (17) possède une seule solution vérifiant les conditions (20). Supposons qu'une telle solution existe. En multipliant (17) par $p(i, k)$ et en sommant sur tous les i, on obtient

$$x_k = \sum_{i\in I} x_i p(i, k) = \sum_{i\in I} \sum_{j\in I} x_j p(j, i) p(i, k) = \\ = \sum_{i\in I} x_j \sum_{i\in I} p(j, i) p(i, k) = \sum_{j\in I} x_j p^{(2)}(j, k).$$

La permutation de l'ordre de sommation est permise par la convergence absolue de la série double. De façon analogue on obtient

$$x_k = \sum_{j\in I} x_j p^{(n)}(j, k). \qquad (22)$$

Posons

$$s_N(j, k) = \frac{1}{N} \sum_{n=1}^{N} p^{(n)}(j, k);$$

alors

$$x_k = \sum_{j\in I} x_j s_N(j, k),$$

$s_N(j, k) \to m_k^{-1}$ et la convergence absolue de la série $\sum_{j\in I} x_j$ donnent par passage à la limite

$$x_k = \sum_{j\in I} x_j m_k^{-1} = m_k^{-1}, \qquad (23)$$

par suite le système (17) possède une solution unique vérifiant (20). Donc si une chaîne est récurrente et nulle, $x_k = 0$, $\forall k \in I$.

Montrons maintenant que dans le cas d'une chaîne positive les quantités (21) forment la solution cherchée du système. Soit I' une partie finie de I. L'inégalité

$$p^{(n+1)}(k, i) \geqslant \sum_{j\in I'} p^{(n)}(k, j) p(j, i)$$

entraîne que

$$s_{N+1}(k, i) - \frac{1}{N+1} p(k, i) \geqslant \frac{N}{N+1} \sum_{j\in I'} s_N(k, j) p(j, i),$$

d'où en faisant tendre $N \to \infty$

$$v_j \geqslant \sum_{j\in I'} v_j p(j, i).$$

Si $I' \to I$, on obtient

$$v_i \geqslant \sum_{j\in I} v_j p(j, i).$$

Une multiplication par $p\ (i,\ k)$ et une sommation sur k nous conduisent aux inégalités

$$v_k \geqslant \sum_{i \in I} v_i p\,(i,\ k) \geqslant \sum_{i \in I} v_i p^{(2)}\,(i,\ k)$$

et, si l'on poursuit cette procédure, aux inégalités

$$v_k \geqslant \sum_{j \in I} v_i p^{(n)}\,(i,\ k),\ \ \forall n \geqslant 1.$$

Avec une inégalité stricte pour au moins un k on aurait eu

$$\sum_{k \in I} v_k > \sum_{j \in I} v_i \sum_{i \in I} p^{(n)}\,(i,\ k) = \sum_{i \in I} v_i,$$

ce qui est impossible. Donc

$$v_k = \sum_{i \in I} v_i p^{(n)}\,(i,\ k),\ \ k \in I,\ \ n = 1,\ 2,\ \ldots \tag{24}$$

En particulier, les v_i constituent la solution du système (17). De (24) on déduit

$$v_l = \sum_{i \in I} v_i s_N\,(i,\ k). \tag{25}$$

On remarquera que l'inégalité

$$\sum_{k \in I'} p^{(n)}\,(i,\ k) \leqslant 1$$

entraîne que

$$\sum_{k \in I'} s_N\,(i,\ k) \leqslant 1 \ \text{ et } \sum_{k \in I'} v_k \leqslant 1,\ \ \forall\, I' \subset I\ (I' \text{ fini}).$$

D'où

$$\sum_{k \in I} v_k \leqslant 1.$$

Donc dans (25) on peut passer à la limite lorsque $N \to \infty$, ce qui donne

$$v_k = \sum_{i \in I} v_i v_k,$$

d'où

$$\sum_{i \in I} v_i = 1,$$

Par suite la solution v_i du système (17) vérifie les conditions (20). ■

Remarque Si la chaîne de Markov est arbitraire, $\{x_i,\ i \in I\}$ est une solution absolument sommable du système (17), et enfin k est un état transitoire, alors $x_k = 0$.

Cette proposition découle de la possibilité de passer à la limite lorsque $n \to \infty$ dans l'égalité (22) et de la relation $\lim\limits_{n \to \infty} p^{(n)}\,(j,\ k) = 0$, qui est réalisée pour tout k transitoire.

COROLLAIRES.

1. *Une condition nécessaire et suffisante pour qu'une chaîne de Markov irréductible soit récurrente positive est que le système* (17) *possède une solution non triviale absolument sommable* $\{x_i,\ i \in I\}$. *De plus* $x_i = cv_i$, *où c est une constante et* $v_i > 0$.

2. *Une condition nécessaire et suffisante pour qu'une chaîne de Markov irréductible possède une répartition initiale invariante est qu'elle soit récurrente positive.*

3. *Si une chaîne est irréductible, positive et apériodique, l'unique solution du système* (17) *vérifiant* (20) *est de la forme*

$$x_i = v_i = \lim_{n\to\infty} p^{(n)}(j,\ i). \tag{26}$$

La dernière proposition découle du fait que $\lim\limits_{n\to\infty} p^{(n)}(j,\ i)$ existent pour une chaîne positive apériodique. Donc (17) résulte de (21).

Le théorème précédent entraîne que dans le cas d'une chaîne récurrente nulle le système (17) ne possède pas de solution non triviale absolument sommable. Mais il admet une solution non sommable non négative importante. Pour l'obtenir on se sert *d'une probabilité de passage avec interdiction ou tabou-probabilité.* Cette notion généralise celle de probabilité de premier accès. On l'a déjà rencontrée dans la démonstration du théorème 13. La *tabou-probabilité* ${}_lp^{(n)}(i,\ j)$ est la probabilité d'arriver en n épreuves à l'état j en partant de l'état initial i sans passer par l'état l aux instants 1, 2, ..., $n - 1$. Donc

$${}_lp^{(n)}(i,\ j) = \sum_{\substack{j_1, j_2, \ldots, j_{n-1},\\ j_r \neq l,\ r=1, \ldots, n-1}} p(i,\ j_1)\, p(j_1,\ j_2) \ \ldots\ p(j_{n-1},\ j),\quad n \geqslant 1.$$

De toute évidence

$${}_lp^{(1)}(i,\ j) = p(i,\ j),\quad {}_jp^{(n)}(i,\ j) = f^{(n)}(i,\ j).$$

Posons encore

$${}_lp^{(0)}(i, j) = \delta(i,\ j).$$

On définit de façon analogue la tabou-probabilité ${}_Hp^{(i,j)}$ attachée à un ensemble interdit d'états H. Si deux états l et j sont interdits, on désignera logiquement la tabou-probabilité ${}_{\{l,j\}}p^{(n)}(i,\ j)$ par ${}_lf^{(n)}(i,\ j)$. Ceci exprime la probabilité d'accéder pour la première fois à l'état j à partir de l'état i en n épreuves sans passer pour cela par l'état l.

Signalons les deux égalités suivantes:

$$_lp^{(n)}(i, j) = \sum_{k=1}^{n} {_lf^{(k)}}(i, j)\, {_lp^{(n-k)}}(j, j), \tag{27}$$

$$_lp^{(n)}(i, j) = \sum_{k=1}^{n} {_lp^{(k)}}(i, i)_{\{i, l\}}\, p^{(n-k)}(i, j). \tag{28}$$

En particulier, la formule (28) entraîne que (pour $l = j$)

$$f^{(n)}(i, j) = \sum_{k=1}^{n} {_jp^{(k)}}(i, i)_i\, f^{(n-k)}(i, j). \tag{29}$$

Soient les fonctions génératrices

$$_lP_{ij}(z) = \sum_{n=0}^{n} {_lp^{(n)}}(i, j)\, z^n,$$

$$_lF_{ij}(z) = \sum_{n=0}^{\infty} {_lf^{(n)}}(i, j)\, z^n, \quad _lf^{(0)}(i, j) = 0.$$

Les seconds membres des égalités (27) et (29) sont l'expression du produit de convolution de deux suites, donc

$$_lP_{ij}(z) = {_lF_{ij}}(z)\, {_lP_{jj}}(z), \quad F_{ij}(z) = {_jP_{ii}}(z)_i\, F_{ij}(z). \tag{30}$$

On remarquera que les séries $_lF_{ij}(z)$ sont convergentes pour $z = 1$ et de plus, si les états i et j sont communicants, alors $_iF_{ij}(1) > 0$. Dans cette hypothèse la deuxième égalité (30) montre qu'il existe une limite finie $_jP_{ii}(z)$ lorsque $z \to 1$ et par suite $_jP_{ii}(1) < \infty$. Posons

$$_lG(i, j) = \sum_{n=0}^{\infty} {_lp^{(n)}}(i, j). \tag{31}$$

Si donc les états i et j sont communicants,

$$_jG(i, i) = \frac{F_{ij}(1)}{_iF_{ij}(1)} < \infty. \tag{32}$$

D'autre part, la première égalité (30) donne

$$_iG(i, j) = {_iF_{ij}}(1)\, {_iG}(j, j),$$

d'où

$$_iG(i, j) \leqslant {_iG}(j, j) < \infty. \tag{33}$$

Revenons au système (17) et prouvons le théorème suivant.

Théorème 15. *Soit l un état arbitraire d'une chaîne de Markov récurrente irréductible. Le système* (17) *possède la solution non négative*

$$x_l = 1, \; x_i = {_lG}(l, i) \; (i \neq l) \; i \in I.$$

D é m o n s t r a t i o n. Posons

$$u_l = 1,\ u_i = {}_lG(l, i)\ (i \neq l). \tag{34}$$

Pour $i \neq l$ on a

$$\sum_{j \in I} u_j p(j, i) = p(l, i) + \sum_{j \neq l} {}_lG(l, j)\, p(j, i) =$$
$$= p(l, i) + \sum_{j \neq l} \sum_{n=1}^{\infty} {}_lp^{(n)}(l, j)\, p(j, i) =$$
$$= p(l, i) + \sum_{n=1}^{\infty} {}_lp^{(n+1)}(l, i) = {}_lG(l, i) = u_i;$$

si $i = l$, alors

$$\sum_{j \in I} u_j p(j, l) = p(l, l) + \sum_{n=1}^{\infty} f^{(n+1)}(l, l) = \sum_{n=1}^{\infty} f^{(n)}(l, l) = 1 = u_l. \ \blacksquare$$

Etudions l'unicité de la solution du système d'équations (17) vérifiant les conditions $u_l = 1$, $u_i \geqslant 0$. Pour ce faire nous allons nous servir d'une méthode faisant intervenir la notion de chaîne de Markov inversée.

Supposons dans un premier temps que la chaîne est récurrente positive et soit $\{v_j, j \in I\}$ une répartition initiale invariante.

Soit $\mathsf{P}^{(v)}$ la mesure probabiliste attachée au processus markovien stationnaire associé à la répartition initiale $\{v_j, j \in I\}$. Soient les probabilités conditionnelles

$$q_i(j_1, j_2, \ldots, j_n) =$$
$$= \mathsf{P}^{(v)}\{\xi(t-1) = j_1,\ \xi(t-2) = j_2, \ldots, \xi(t-n) =$$
$$= j_n \mid \xi(t) = i\},$$

où $t > n$; on a

$$q_i(j_1, j_2, \ldots, j_n) = \frac{v_{j_n} p(j_n, j_{n-1})\, p(j_{n-1}, j_{n-2}) \ldots p(j_1, i)}{v_i} =$$
$$= q(i, j_1)\, q(j_1, j_2) \ldots q(j_{n-1}, j_n),$$

où $q(i, j) = p(j, i) \dfrac{v_j}{v_i}$.

Donc, dans une chaîne de Markov récurrente positive stationnaire, les probabilités conditionnelles de passage résultant d'une inversion du temps (c'est-à-dire le temps est compté du présent vers le passé) sont également attachées à une chaîne de Markov. Cela étant, $v_i > 0$ entraîne que

$$q(i, j) \geqslant 0, \quad \sum_{j \in I} q(i, j) = \frac{1}{v_i} \sum_{j \in I} v_j p(j, i) = \frac{v_i}{v_i} = 1.$$

Cette construction peut être effectuée aussi bien pour les chaînes récurrentes positives que pour les récurrentes nulles. Considérons à

ces fins une solution positive quelconque $\{x_j, j \in I\}$ du système (17) (l'existence d'une telle solution sera prouvée plus bas) et posons

$$q(i, j) = p(j, i)\frac{x_j}{x_i}. \tag{35}$$

Comme précédemment,

$$q(i, j) \geqslant 0, \sum_{j \in I} q(i, j) = 1.$$

Une chaîne de Markov de probabilités de passage (35) sera appelée *chaîne inverse* de l'initiale.

Signalons les formules des probabilités de passage en n épreuves dans une chaîne inverse. On a

$$q^{(n)}(i, j) = \sum_{j_1, j_2, \ldots, j_{n-1}} q(i, j_1) q(j_1, j_2) \ldots q(j_{n-1}, j) =$$

$$= \sum_{j_1, j_2, \ldots, j_{n-1}} p(j_1, i) p(j_2, j_1) \ldots p(j, j_{n-1}) \frac{x_j}{x_i},$$

c'est-à-dire

$$q^{(n)}(i, j) = \frac{x_j}{x_i} p^{(n)}(j, i). \tag{36}$$

Par suite, si la chaîne initiale est irréductible, récurrente, positive ou nulle, il en sera de même pour la chaîne inverse.

L'application de la relation (10) à une chaîne inverse donne

$$\lim_{N \to \infty} \frac{\sum_{n=1}^{N} q^{(n)}(i, j)}{\sum_{n=0}^{N} q^{(n)}(j, j)} = 1.$$

En vertu des formules (36) on a

$$\lim_{N \to \infty} \frac{\sum_{n=1}^{N} p^{(n)}(j, i)}{\sum_{n=0}^{N} p^{(n)}(j, j)} = \frac{x_i}{x_j}, \tag{37}$$

d'où le

THÉORÈME 16. *Dans le cas d'une chaîne récurrente irréductible le système* (17) *admet une solution non négative unique telle que* $x_l = 1$. *De plus* $x_i = {}_lG(l, i)$ *et*

$$\lim_{N \to \infty} \frac{\sum_{n=1}^{N} p^{(n)}(j, i)}{\sum_{n=0}^{N} p^{(n)}(j, j)} = {}_jG(j, i). \tag{38}$$

La formule (38) découle de l'unicité de la solution du système (17) et du théorème (15), l'unicité, de la formule (37) sous réserve que $x_j > 0$, $\forall j$. Donc d'après le théorème 15 il suffit de prouver que si $\{x_j,\ j \in I\}$ est solution non triviale non négative du système (17), alors $x_j > 0$. On montre que $x_j > 0$ de la façon suivante. Pour toute solution non négative du système (17) on a

$$x_i = \sum_j x_j p(j, i) = \sum_j \sum_k x_k p(k, j)\, p(j, i) =$$
$$= \sum_k x_k \sum_j p(k, j)\, p(j, i) = \sum_k x_k p^{(2)}(k, i).$$

Par récurrence on démontre sans peine que

$$x_i = \sum_{k \in I} x_k p^{(n)}(k, i).$$

Soit $x_l > 0$;

$$\forall i\ \exists n \text{ tel que } p^{(n)}(l, i) > 0;$$

donc

$$x_i \geqslant x_l p^{(n)}(l, i) > 0.$$

En construisant pour cette solution une chaîne inverse et en faisant $x_l = 1$, on déduit de (37) l'unicité de x_i, $i \in I$. Le théorème 15 donne $x_i = {}_iG(l, i)$. ■

Remarque. La formule (37) généralise la relation

$$\lim_{N\to\infty} N^{-1} \sum_{n=1}^{N} p^{(n)}(j, i) = v_i,$$

où $\{v_i\}$ est une répartition initiale invariante d'une chaîne récurrente positive irréductible.

Théorème 17. *Dans le cas d'une chaîne de Markov récurrente irréductible le système d'inéquations*

$$x_i \geqslant \sum_{j \in I} x_j p(j, i), \quad x_i \geqslant 0, \ x_l = 1, \tag{39}$$

possède une solution unique et de plus

$$x_i = \sum x_j p(j, i), \quad i \in I.$$

D'après le théorème 16 il suffit simplement de prouver l'unicité de la solution du système (39). Soit une chaîne de Markov inverse de probabilités de passage

$$q(i, j) = p(j, i)\frac{u_j}{u_i},$$

où u_i est une solution positive du système (17). C'est une chaîne irréductible et récurrente. On a

$$\sum_j q(i, j) \frac{x_j}{u_j} = \sum_j p(j, i) \frac{x_j}{u_i} \leqslant \frac{x_i}{u_i}, \quad \frac{x_l}{u_l} = 1.$$

Or, d'après le théorème 13, le système d'inéquations

$$\sum_j q(i, j) y_j \leqslant y_i, \quad y_l = 1,$$

possède une solution non négative unique $y_i = 1$. Donc

$$x_i = u_i \ \forall i \in I. \ \blacksquare$$

CHAPITRE IV

FONCTIONS ALÉATOIRES

§ 1. Définition d'une fonction aléatoire

Au chapitre I la fonction aléatoire a été définie comme une famille de variables aléatoires dépendant d'un paramètre. On y a mentionné les difficultés liées à cette définition au sens large, c'est-à-dire une collection de fonctions de répartition finidimensionnelles vérifiant des conditions de compatibilité. L'axiomatique de la théorie des probabilités nous suggère d'entendre naturellement par fonction aléatoire une famille arbitraire de variables aléatoires définies sur un même espace probabilisé.

DÉFINITION. *Soit* $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ *un espace probabilisé. On appelle fonction aléatoire une fonction de deux variables* $g(\theta, \omega) = \xi(\theta)$ *définie pour* $\theta \in \Theta$, $\omega \in \Omega$, *à valeurs dans un espace métrique* X, $\mathfrak{S}$*-mesurable comme fonction de* ω *pour chaque* $\theta \in \Theta$. *L'ensemble* Θ *est le domaine de définition de la fonction aléatoire,* X *le domaine de valeurs.*

Le cas particulier suivant de la définition générale est intéressant. Supposons que Ω est un espace fonctionnel, $\omega = \omega(\theta)$, $\theta \in \Theta$, et que la tribu $\mathfrak{S}$ contient tous les ensembles de Ω de la forme

$$\{\omega\colon \omega(\theta) \in A\}$$

quel que soit $\theta \in \Theta$ et quel que soit l'ensemble borélien $A \in X$, P est une mesure probabiliste arbitraire sur $\mathfrak{S}$. Il semble naturel d'assimiler un tel espace probabilisé à une fonction aléatoire $g(\theta, \omega) = \omega(\theta)$. Parfois il est commode d'identifier la fonction aléatoire $g(\theta, \omega) = \omega(\theta)$ à un espace probabilisé du type décrit.

On remarque aisément que la définition générale de la fonction aléatoire peut être ramenée au cas particulier décrit. En effet, si une fonction aléatoire $\xi(\theta)$ est donnée comme une fonction de deux variables $\xi(\theta) = g(\theta, \omega)$, alors en posant $u = g(\theta, \omega)$, où ω est fixe, $\omega \in \Omega$, et en désignant par U l'ensemble de toutes les fonctions $u = g(\theta, \omega)$ obtenues lorsque ω parcourt Ω, on obtient une application T de Ω sur U qui associe à la tribu $\mathfrak{S}$ des ensembles de Ω une tribu $\mathfrak{F}$ d'ensembles de U et à la mesure probabiliste P sur $\mathfrak{S}$ une mesure probabiliste P' sur $\mathfrak{F}$ (cf. § 6, chap. II). Pour tout θ fixe

14—0398

l'ensemble $\{u : u = g(\theta, \omega) < x\}$ appartient à $\mathfrak{F}$, puisque $T^{-1}\{u : u = g(\theta, \omega) < x\} = \{\omega : g(\theta, \omega) < x\} \in \mathfrak{S}$.

On obtient donc un espace probabilisé $\{U, \mathfrak{F}, \mathsf{P}'\}$, où U est un ensemble de fonctions $u = u(\theta)$ tel que, quels que soient n, $\theta_1, \theta_2, \ldots$ $\ldots, \theta_n$ ($\theta_k \in \Theta$, $k = 1, \ldots, n$), la répartition de la suite de variables aléatoires

$$g(\theta_1, \omega),\ g(\theta_2, \omega),\ \ldots,\ g(\theta_n, \omega)$$

sur $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ coïncide avec celle de la suite de variables

$$u(\theta_1),\ u(\theta_2),\ \ldots,\ u(\theta_n)$$

définies sur l'espace probabilisé $\{U, \mathfrak{F}, \mathsf{P}'\}$.

Enonçons maintenant un point de vue, capital pour la suite, sur l'équivalence de fonctions aléatoires. Dans les problèmes il n'y a aucune raison de différencier des fonctions aléatoires se déduisant l'une de l'autre par des transformations de l'espace probabilisé. Bien plus, d'un point de vue pratique, l'expérience ne permet de distinguer que les hypothèses relatives aux répartitions finidimensionnelles d'une fonction aléatoire. Aussi admet-on que les données empiriques ne permettent pas de faire une distinction entre deux fonctions aléatoires $\xi(\theta)$ et $\xi'(\theta)$ dont coïncident les répartitions finidimensionnelles, c'est-à-dire les répartitions conjointes des suites

$$\xi(\theta_1),\ \xi(\theta_2),\ \ldots,\ \xi(\theta_n) \tag{1}$$

et

$$\xi'(\theta_1),\ \xi'(\theta_2),\ \ldots,\ \xi'(\theta_n) \tag{2}$$

pour tous les entiers $n \geqslant 1$ et $\theta_k \in \Theta$, $k = 1, \ldots, n$. On adoptera donc la définition suivante.

DÉFINITION. *On dit que deux fonctions aléatoires* $\xi(\theta)$ *et* $\xi'(\theta)$ *ayant le même domaine de définition* Θ *sont stochastiquement équivalentes au sens large si, quels que soient* $n \geqslant 1$ *entier et* $\theta_k \in \Theta$, $k = 1, 2, \ldots, n$, *les répartitions conjointes des suites de variables aléatoires* (1) *et* (2) *sont confondues.*

Dans la suite il sera fait un large usage des fonctions aléatoires stochastiquement équivalentes dans un sens plus étroit.

DÉFINITION. *Deux fonctions aléatoires* $g_1(\theta, \omega)$ *et* $g_2(\theta, \omega)$ ($\theta \in \Theta$, $\omega \in \Omega$) *définies sur un même espace probabilisé sont stochastiquement équivalentes si pour tout* $\theta \in \Theta$

$$\mathsf{P}\{g_1(\theta, \omega) \neq g_2(\theta, \omega)\} = 0.$$

Il est évident que si $g_1(\theta, \omega)$ et $g_2(\theta, \omega)$ sont stochastiquement équivalentes, elles le seront au sens large.

Voyons quelques exemples de fonctions aléatoires.

a) L'oscillation aléatoire $\zeta(t)$, $(-\infty < t < \infty)$ étudiée au § 5, chapitre I,

$$\zeta(t) = \sum_{k=1}^{n} \gamma_k e^{iu_k t}, \quad \gamma_k = \alpha_k + i\beta_k,$$

α_k, β_k $(k = 1, 2, \ldots, n)$ étant des variables aléatoires, peut être mise sous la forme

$$\zeta(t) = g(t, \omega),$$

où $\omega = (\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_n)$ est un point d'un espace réel $\mathcal{R}^{2n}$ $2n$-dimensionnel et $g(t, \omega)$ une fonction linéaire de ω à t fixe et une somme de fonctions trigonométriques de t à ω fixe. Dans $\mathcal{R}^{2n}$ la probabilité P est donnée par la répartition conjointe des variables aléatoires $\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_n$. D'autre part, le processus $\zeta(t)$ peut être traité comme un espace probabilisé $\{U, \mathfrak{S}, \mathsf{P}'\}$, où U est l'espace de toutes les fonctions à valeurs complexes de la forme $u = \sum_{k=1}^{n} \gamma_k e^{iu_k t}$, la base des exposants $(u_1, u_2, \ldots, u_n)$ étant donnée.

La mesure P' sur U est induite par la mesure P par l'application $\omega \to u = g(t, \omega)$.

b) Soit un processus $\xi(t)$ dont les réalisations sont constantes sur les intervalles de temps $[k-1, k]$ et y prennent la valeur 0 ou 1 avec des probabilités identiques ne dépendant pas des valeurs prises sur les intervalles de temps antérieurs.

Entre les réalisations du processus $\xi(t)$ et les fractions binaires infinies on peut établir une correspondance bi-univoque:

$$\xi(t) \to \omega = 0, x_1 x_2 x_3 \ldots x_n \ldots, \tag{3}$$

où x_n est la valeur prise par $\xi(t)$ sur l'intervalle $[n-1, n]$. A toute fraction binaire infinie ω est associé un point de l'intervalle $[0, 1]$. La bi-univocité de cette correspondance est violée seulement pour les fractions dont tous les chiffres binaires sont confondus à partir d'un certain rang. A deux telles fractions distinctes correspond un même point de l'intervalle $[0, 1]$ (exception faite des points 0 et 1 dont l'unique représentation sous forme d'une fraction binaire infinie est: $0 = 0, 00 \ldots 0 \ldots$, $1 = 0, 11 \ldots 1 \ldots$). Les réalisations correspondant aux fractions binaires infinies deviennent constantes à partir d'un certain instant. Désignons par $\mathfrak{A}$ l'ensemble de ces réalisations et par $\mathfrak{A}_n$ l'événement qui consiste en ce que la fonction $\xi(t)$ est constante à partir d'un instant n. On a

$$\mathfrak{A} = \bigcup_{n=1}^{\infty} \mathfrak{A}_n.$$

Supposons encore que $\mathfrak{A}_{nm}$ signifie que la fonction $\xi(t)$ est constante sur l'intervalle $[n, n+m]$. Les événements $\mathfrak{A}_{nm}$ $(m = 1, 2, \ldots)$

forment une suite monotone décroissante, $\mathfrak{A}_n = \bigcap_{m=1}^{\infty} \mathfrak{A}_{nm}$ et

$$\mathsf{P}(\mathfrak{A}_n) = \lim_{m\to\infty} \mathsf{P}(\mathfrak{A}_{nm}).$$

$$\mathsf{P}(\mathfrak{A}_{nm}) = \frac{1}{2^{m-1}} \Rightarrow \mathsf{P}(\mathfrak{A}_n) = 0 \text{ et } \mathsf{P}(\mathfrak{A}) = 0.$$

Par suite, si l'on néglige les fonctions aléatoires $\xi(t)$ de probabilité nulle, on établit une correspondance bi-univoque entre les points de l'intervalle $[0, 1]$ et les fonctions aléatoires $\xi(t)$.

Soit Δ un intervalle de longueur $1/2^n$ séparé par des nombres binaires rationnels, c'est-à-dire un intervalle de la forme $[0, j_1 \ldots j_n;$ $0, j_1 \ldots j_n + 1[$, où j_k prennent la valeur 0 ou 1, $k = 1, 2, \ldots, n$.

L'ensemble A des fonctions aléatoires $\xi(t)$ associées aux points $\omega \in \Delta$ est composé de fonctions vérifiant les conditions

$$\xi(0) = j_1, \ \xi(1) = j_2, \ \ldots, \ \xi(n-1) = j_n.$$

La probabilité que $\xi(t) \in A$ est égale à $1/2^n$, c'est-à-dire à la longueur de l'intervalle Δ. D'où il suit que si B est un borélien quelconque de points de l'intervalle $[0, 1]$ et B' l'ensemble des fonctions $\xi(t)$ mis par (3) en relation bi-univoque avec l'ensemble numérique B, alors $\mathsf{P}(B')$ est confondue avec la mesure de Lebesgue de B. Donc choisir une fonction aléatoire $\xi(t)$ revient à choisir arbitrairement un point ω dans $[0, 1]$ muni de la mesure de Lebesgue. Plus exactement, cela veut dire que le processus aléatoire se décrit entièrement de la façon suivante. On prend un point quelconque de l'intervalle $[0, 1]$. La probabilité que $\omega \in B$, où B est un borélien de $[0, 1]$, est égale à la mesure de Lebesgue de B. L'écriture binaire de la coordonnée du point ω est $0, x_1 x_2 \ldots x_n \ldots$ La fonction $\xi(t)$ prend alors la valeur x_n sur l'intervalle $[(n-1), n]$. Donc $\xi(t)$ s'écrit encore

$$\xi(t) = f(t, \omega),$$

où $f(t, \omega)$ est une fonction non aléatoire, complètement définie de $t \in [0, \infty[$ et $\omega \in [0, 1]$.

c) Soit une fonction aléatoire au sens large à valeurs dans X, espace polonais (métrique, complet et séparable), et définie sur Θ. On a vu au § 2, chapitre II, qu'on pouvait toujours construire un espace probabilisé $\{\Omega, \tilde{\mathfrak{S}}, \mathsf{P}\}$, où Ω est l'ensemble de toutes les applications $\omega = \omega(\theta)$ de l'ensemble Θ dans X, tel que la répartition des suites $\{\omega(\theta_1), \ldots, \omega(\theta_n)\}$ de X^n quels que soient n, $\theta_k \in \Theta$, $k = 1, 2, \ldots, n$, soit confondue avec celle de la fonction aléatoire au sens large. Autrement dit, on peut toujours construire une fonction aléatoire stochastiquement équivalente au sens large (et représentée dans les termes du § 2, chapitre II) à une fonction aléatoire au sens large.

Malheureusement, la représentation d'une fonction aléatoire, préconisée par le théorème de Kolmogorov, n'est pas judicieuse. Les événements élémentaires de l'espace probabilisé construit sont des fonctions arbitraires $\omega = \omega(\theta)$ et il ne devient plus possible d'envisager leur continuité, intégrabilité, dérivabilité, etc. (il est entendu que ces notions ont un sens dans Θ et X).

D'autre part, l'espace probabilisé construit ne contient pas les événements, importants pour la résolution de nombreux problèmes, de la forme

$$\{\omega(\theta) \in A, \ \forall \theta \in Q\}, \tag{4}$$

où Q est partie non dénombrable de Θ et $A \subset X$. En effet, à l'événement

$$\{\omega : \omega = \omega(\theta) \in A, \ \forall \theta \in Q\} = \bigcap_{\theta \in Q} \{\omega : \omega(\theta) \in A\}$$

est associé dans Ω un ensemble intersection d'un nombre non dénombrable d'ensembles de $\tilde{\mathfrak{S}}$ et par suite il n'est pas tenu d'appartenir à la tribu $\tilde{\mathfrak{S}}$. Ceci nous suggère de prendre un espace Ω aussi étroit que possible, c'est-à-dire un espace dont les fonctions posséderaient de meilleures propriétés analytiques.

Par exemple, pour calculer la probabilité de l'événement (4) il est souhaitable que la fonction aléatoire $g(\theta, \omega)$ qui est la représentation de la famille de répartitions (1) possède la propriété suivante:

p) Pour des classes assez vastes d'ensembles $\mathfrak{A}$ de X et $\mathfrak{Q}$ de Θ il existe un ensemble dénombrable S de points $\theta_j \in \Theta$ tel que, quels que soient $A \in \mathfrak{A}$ et $Q \in \mathfrak{Q}$, l'ensemble des points

$$\{\omega: g(\theta, \omega) \in A \ \forall \ \theta \in Q\}, \ A \in \mathfrak{A}, \ Q \in \mathfrak{Q}, \tag{5}$$

coïncide avec l'ensemble

$$\{\omega: g(\theta_j, \omega) \in A \ \forall \ \theta_j \in S \cap Q\} \tag{6}$$

à une partie près d'un ensemble fixe N, de $\mathbf{P}$-mesure nulle, ne dépendant ni de A ni de Q.

Etant intersection d'une suite au plus dénombrable d'ensembles mesurables, l'ensemble (6) est lui-même mesurable de même que l'ensemble (5) (en vertu de la complétude de la mesure $\mathbf{P}$). Cela étant, les événements (5) et (6) ont même probabilité.

Toute fonction aléatoire possédant la propriété p) est dite *séparable* (par rapport à la classe d'ensembles $\mathfrak{A}$).

On remarquera que si Θ est un espace métrique séparable, $\mathfrak{Q}$ une classe d'ouverts, $\mathfrak{A}$ une classe de fermés de X et $g(\theta, \omega)$ une fonction continue en θ à ω fixe pour presque tous les ω, alors $g(\theta, \omega)$ est séparable par rapport à la classe de fermés $\mathfrak{A}$.

Dans les problèmes faisant intervenir l'intégration de variables aléatoires par rapport à une mesure $\{m, \mathfrak{K}\}$ sur Θ, il est souhaitable

que la fonction $g(\theta, \omega)$ soit $\mathfrak{K}$-mesurable comme fonction de θ à ω fixe P-presque pour tous les ω.

Dans d'autres cas il faut trouver une représentation de la fonction donnée au sens large, telle que presque toutes les réalisations soient continues ou bien présentent seulement des discontinuités de première espèce ou bien soient k fois dérivables, etc. De toute évidence la possibilité d'obtenir une représentation possédant des propriétés spéciales dépend probablement des répartitions finidimensionnelles de la fonction aléatoire.

Ces problèmes ne se posent pas uniquement pour les fonctions aléatoires au sens large. La fonction aléatoire $g(\theta, \omega)$ est susceptible de posséder des « propriétés pathologiques », mais elle peut très bien avoir une fonction « lissée » $g^*(\theta, \omega)$, $\mathsf{P}\{g(\theta, \omega) \neq g^*(\theta, \omega)\} = 0$ qui lui est stochastiquement équivalente et qui n'est plus pathologique. Le point de vue adopté nous permet de remplacer une telle fonction par une fonction régulière $g^*(\theta, \omega)$ stochastiquement équivalente.

EXEMPLE. Soient A l'ensemble des rationnels de la droite $(-\infty, \infty)$, $\chi(t)$ l'indicateur de A et ω une variable aléatoire équirépartie sur $[0, 1]$.

Posons $g(t, \omega) = \chi(t + \omega)$. La fonction $g(t, \omega)$ est partout discontinue à ω fixe. D'autre part, elle est presque sûrement nulle à t fixe. Par suite elle est stochastiquement équivalente à la fonction $g^*(t, \omega) \equiv 0$.

Dans ce paragraphe on se propose d'étudier les conditions d'existence d'une fonction possédant certaines propriétés de régularité et stochastiquement équivalente (ou stochastiquement équivalente au sens large) à la fonction donnée.

§ 2. Fonctions aléatoires séparables

La notion de fonction aléatoire séparable a déjà été introduite au § 1. Il se trouve que la propriété de séparabilité n'est pas une restriction forte sur la fonction aléatoire. Sous des hypothèses assez larges concernant uniquement la nature du domaine de définition Θ et du domaine de valeurs X de la fonction aléatoire, il existe une fonction aléatoire séparable stochastiquement équivalente à la fonction donnée. A noter toutefois que lorsqu'on construit une fonction aléatoire séparable équivalente, force est parfois d'élargir le domaine des valeurs de la fonction en le transformant en un ensemble compact.

Dans le présent paragraphe on admet constamment que Θ et X sont des espaces métriques munis des distances $r(\theta_1, \theta_2)$ et $\rho(x_1, x_2)$ respectivement et que Θ est séparable. Les classes d'ensembles $\mathfrak{A}$ et $\mathfrak{D}$ intervenant dans la définition de la séparabilité sont des

fermés de X et des ouverts de Θ. Donc la séparabilité d'une fonction aléatoire est comprise au sens suivant.

DÉFINITION. *Une fonction aléatoire* $g(\theta, \omega)$ *est séparable si existent dans* Θ *un ensemble partout dense de points* $\{\theta_j\}$, $j = 1, 2, \ldots$, *et dans* Ω *un ensemble* N *de probabilité nulle, tels que pour tout ouvert* $G \subset \Theta$ *et tout fermé* $F \subset X$ *les deux ensembles*

$$\{\omega : g(\theta_j, \omega) \in F,\ \theta_j \in G\},$$
$$\{\omega : g(\theta, \omega) \in F,\ \forall \theta \in G\}$$

coïncident à une partie de N *près.*

L'ensemble dénombrable de points $\{\theta_j\}$ est appelé *ensemble de séparabilité* de la fonction aléatoire.

THÉORÈME 1. *Soient* X *et* Θ *deux espaces métriques,* X *compact,* Θ *séparable. Toute fonction aléatoire* $g(\theta, \omega)$ *définie sur* Θ *et à valeurs dans* X *est stochastiquement équivalente à une fonction aléatoire séparable.*

Démontrons préalablement trois lemmes. Soient $\tilde{g}(\theta, \omega)$ une fonction aléatoire séparable, I son ensemble de séparabilité et N l'espace exclusif correspondant de points ω.

Nommons V la classe de toutes les boules ouvertes de l'espace Θ, de rayons rationnels, centrées aux points d'un ensemble fixe dénombrable partout dense dans Θ. La classe V est dénombrable. Par ailleurs, tout ouvert $G \subset \Theta$ est représentable par la somme (d'un nombre dénombrable) de boules de V.

Supposons que $A(G, \omega)$ est la fermeture de l'ensemble de valeurs de la fonction $\tilde{g}(\theta, \omega)$ lorsque θ parcourt l'ensemble $I \cap G$ et que

$$A(\theta, \omega) = \bigcap_{S \in V} A(S, \omega)$$

est l'intersection de tous les $A(S, \omega)$ lorsque S parcourt la collection des boules contenant le point θ. La famille d'ensembles fermés $A(S, \omega)$ $(\theta \in S)$ est centrée, c'est-à-dire tout nombre fini d'ensembles de cette famille possèdent des points en commun, et en vertu de la compacité de X leur intersection $A(\theta, \omega)$ n'est pas vide. Par ailleurs $A(\theta, \omega)$ est fermé. La séparabilité de la fonction $\tilde{g}(\theta, \omega)$ implique (pour $\omega \notin N$)

$$\tilde{g}(\theta, \omega) \in A(\theta, \omega). \tag{1}$$

Inversement, si (1) est réalisée pour tout $\omega \notin N$, $\mathsf{P}(N) = 0$, alors $\tilde{g}(\theta, \omega)$ est une fonction aléatoire séparable. En effet, si $\tilde{g}(\theta, \omega) \in F$ pour tous les $\theta \in I \cap S$, où F est un fermé de X et $S \in V$, alors $A(\theta, \omega) \subset A(S, \omega) \subset F$ pour tout $\theta \in S$ et par suite $\tilde{g}(\theta, \omega) \in F$ pour tous les $\theta \in S$.

Si G est un ouvert de Θ, il suffit de le représenter par la somme $G = \bigcup_k S_k$ d'ensembles de V pour s'assurer d'après ce qui précède que

$$\tilde{g}(\theta, \omega) \in F \quad \forall \theta \in I \cap G, \quad \omega \notin N,$$

entraîne

$$g(\theta, \omega) \in F \quad \forall \theta \in G.$$

D'où le

LEMME 1. *Pour qu'une fonction aléatoire* $\tilde{g}(\theta, \omega)$ *soit séparable il faut et il suffit qu'existe un ensemble* N *de probabilité nulle tel que* (1) *soit réalisé pour* $\omega \notin N$.

Donc pour construire une fonction séparable stochastiquement équivalente à $g(\theta, \omega)$ il suffit de trouver une fonction $\tilde{g}(\theta, \omega)$ vérifiant (1) et telle que pour tout $\theta \in \Theta$ on ait

$$\mathsf{P}\{\tilde{g}(\theta, \omega) \neq g(\theta, \omega)\} = 0.$$

LEMME 2. *Soit* B *un borélien de* X. *Il existe une suite finie ou dénombrable de points* $\theta_1, \theta_2, \ldots$ *telle que l'ensemble* $N(\theta, B) = \{\omega : g(\theta_k, \omega) \in B,\ k = 1, 2, \ldots, g(\theta, \omega) \notin B\}$ *est de probabilité nulle pour tout* $\theta \in \Theta$.

D é m o n s t r a t i o n. Soit θ_1 quelconque. Si la suite $\theta_1, \theta_2, \ldots, \theta_k$ est construite, on pose

$$m_k = \sup_{\theta \in \Theta} \mathsf{P}\{g(\theta_1, \omega) \in B, \ldots, g(\theta_k, \omega) \in B;\quad g(\theta, \omega) \notin B\}.$$

La suite m_k est monotone décroissante. Si $m_k = 0$, alors la suite cherchée est construite. Si $m_k > 0$, soit θ_{k+1} un point tel que

$$\mathsf{P}\{g(\theta_1, \omega) \in B, \ldots, g(\theta_k, \omega) \in B,\ g(\theta_{k+1}, \omega) \notin B\} \geqslant \frac{m_k}{2}.$$

Les ensembles

$$L_k = \{\omega : g(\theta_i, \omega) \in B,\ i = 1, 2, \ldots, k,\ g(\theta_{k+1}, \omega) \notin B\}$$

étant disjoints,

$$1 \geqslant \sum_{k=1}^{\infty} \mathsf{P}(L_k) \geqslant \frac{1}{2} \sum_{k=1}^{\infty} m_k.$$

Par suite, $m_k \to 0$ lorsque $k \to \infty$ et quel que soit θ

$$\mathsf{P}\{g(\theta_k, \omega) \in B,\ k = 1, 2, \ldots, g(\theta, \omega) \notin B\} \leqslant \lim m_k = 0. \ \blacksquare$$

Du lemme 2 on déduit sans peine le

LEMME 3. *Soient* $\mathfrak{M}_0$ *une classe dénombrable d'ensembles,* $\mathfrak{M}$ *la classe constituée des intersections de toutes les suites d'ensembles de* $\mathfrak{M}_0$. *Il existe une suite finie ou dénombrable de points* $\theta_1, \theta_2, \ldots, \theta_n, \ldots$ *et pour tout* θ *un ensemble* $N(\theta)$ *tel que*

$$\mathsf{P}\{N(\theta)\} = 0$$

et

$$\{\omega : g(\theta_n, \omega) \in B,\ n = 1, 2, \ldots,\ g(\theta, \omega) \notin B\} \subset N(\theta)$$

$\forall B \in \mathfrak{M}$.

D é m o n s t r a t i o n. Soient I un ensemble dénombrable de points de Θ, somme des suites $\{\theta_n, n = 1, 2, \ldots\}$ construites pour tout $B \in \mathfrak{M}_0$ suivant le lemme 2, et $N(\theta) = \bigcup_{B \in \mathfrak{M}_0} N(\theta, B)$. Si

$$B' \in \mathfrak{M},\ B \supset B',\quad B \in \mathfrak{M}_0,$$

alors

$$\{\omega : g(\theta_n, \omega) \in B',\ \theta_n \in I,\ g(\theta, \omega) \notin B\} \subset$$
$$\subset \{\omega : g(\theta_n, \omega) \in B,\ \theta_n \in I,\ g(\theta, \omega) \notin B\} \subset$$
$$\subset N(\theta, B) \subset N(\theta).$$

D'autre part, si

$$B' = \bigcap_{k=1}^{\infty} B_k,\ B_k \in \mathfrak{M}_0,$$

alors

$$\{\omega : g(\theta_n, \omega) \in B',\ \theta_n \in I,\ g(\theta, \omega) \notin B'\} \subset$$
$$\subset \bigcup_{k=1}^{\infty} \{\omega : g(\theta_n, \omega) \in B',\ \theta_n \in I,\ g(\theta, \omega) \notin B_k\} \subset$$
$$\subset \bigcup_{k=1}^{\infty} N(\theta, B_k) \subset N(\theta),$$

d'où le lemme. ■

La démonstration du théorème 1 est désormais chose simple.

Fixons un ensemble L de points dénombrables partout dense dans X et appelons $\mathfrak{M}_0$ la classe des complémentaires des boules de rayons rationnels centrées en des points de L. Alors $\mathfrak{M}$, la classe des intersections des ensembles de $\mathfrak{M}_0$, contient tous les fermés. Pour chaque $S \in V$, traitons la fonction aléatoire $g(\theta, \omega)$ comme donnée uniquement pour $\theta \in S$, puis formons la suite $I = I(S)$ et les ensembles $N(\theta) = N_S(\theta)$ suivant le lemme 3. Soit

$$J = \bigcup_{S \in V} I(S),\quad N_\theta = \bigcup_{S \in V} N_S(\theta).$$

Posons

$$\tilde{g}(\theta, \omega) = g(\theta, \omega) \text{ si } \theta \in J \text{ ou } g(\theta, \omega) \bar{\in} N_\theta;$$

si $g(\theta, \omega) \in N_\theta$, $\theta \notin J$, on définira $\tilde{g}(\theta, \omega)$ d'une façon quelconque mais de telle sorte que $\tilde{g}(\theta, \omega) \in A(\theta, \omega)$. Les fonctions $\tilde{g}(\theta, \omega)$ et $g(\theta, \omega)$ prenant les mêmes valeurs aux points $\theta \in J$, les ensembles $A(\theta, \omega)$ construits pour les fonctions $\tilde{g}(\theta, \omega)$ et $g(\theta, \omega)$ sont con-

fondus. La définition implique

$$\tilde{g}(\theta,\ \omega) \in A(\theta,\ \omega)$$

pour tous les θ et ω.

$$\{\omega : g(\theta,\ \omega) \neq \tilde{g}(\theta, \omega)\} \subset N_\theta \Rightarrow \mathsf{P}\ \{\tilde{g}(\theta,\ \omega) = g(\theta,\ \omega)\} = 1. \ \blacksquare$$

Le théorème 1 se généralise immédiatement aux fonctions à valeurs dans des espaces localement compacts séparables.

THÉORÈME 2. *Soient X espace séparable localement compact et Θ espace séparable métrique. Toute fonction aléatoire $g(\theta,\ \omega)$ définie sur Θ et à valeurs dans X possède une fonction aléatoire séparable stochastiquement équivalente $\tilde{g}(\theta,\ \omega)$ à valeurs sur une extension compacte $\tilde{X}$ de l'espace X, $\tilde{X} \supset X$.*

La démonstration découle du fait que tout espace séparable localement compact X peut être traité comme un sous-ensemble d'un compact $\tilde{X}$. Par exemple, si $g(\theta,\ \omega)$ est une fonction aléatoire à valeurs dans un espace finidimensionnel X, en ajoutant à X le point ∞, on obtient sans peine un espace compact $X = X \cup \{\infty\}$ muni d'une nouvelle métrique et tel que tout fermé $F \subset X$ (dans la topologie de X) est un fermé dans $\tilde{X}$ (par rapport à la nouvelle métrique). Lorsqu'on construira une réalisation séparable de la fonction aléatoire, il faudra probablement lui attribuer la valeur « ∞ »; mais de toute évidence, si θ est fixe, la probabilité de cette circonstance est nulle. ■

Dans beaucoup de problèmes il est important de savoir quel ensemble J est susceptible d'être ensemble de séparabilité.

THÉORÈME 3. *Soient Θ un espace séparable et $g(\theta,\ \omega)$ une fonction aléatoire séparable stochastiquement continue. Tout ensemble de points de Θ, dénombrable partout dense, est susceptible d'être ensemble de séparabilité de la fonction aléatoire $g(\theta,\ \omega)$.*

D é m o n s t r a t i o n. Soient $V = \{S\}$ l'ensemble dénombrable de boules de Θ considéré dans la démonstration du théorème 1, $J = \{\theta_k,\ k = 1,\ 2,\ \ldots,\ n,\ \ldots\}$ un ensemble de séparabilité de la fonction aléatoire $g(\theta,\ \omega)$, N l'ensemble exclusif des valeurs de ω qui ont servi à définir la séparabilité, Λ un ensemble de points partout dense dans Θ. Soient $B(S,\ \omega)$ la fermeture de l'ensemble des valeurs de $g(\theta,\ \omega)$ lorsque le point θ parcourt $\Lambda \cap S$ et $N(S,\ k)$ l'événement : $g(\theta_k,\ \omega) \bar{\in} B(S,\ \omega)$ si $\theta_k \in S$. Les événements $N(S,\ k)$ sont de probabilité nulle. En effet, soit γ_r. $r = 1,\ 2,\ \ldots,\ n,\ \ldots$, une suite arbitraire de points de $\Lambda \cap S$ convergeant vers θ_k. Alors

$$\mathsf{P}\{g(\theta_k,\ \omega) \notin B(S,\ \omega)\} \leqslant \mathsf{P}\{\varliminf_{r\to\infty} \rho(g(\theta_k,\ \omega),\ g(\gamma_r,\ \omega)) > 0\} \leqslant$$

$$\leqslant \varliminf_{n\to\infty} \mathbf{P}\left\{\varliminf_{r\to\infty} \rho\left(g(\theta_k, \omega), g(\gamma_r, \omega)\right) > \frac{1}{n}\right\} \leqslant$$

$$\leqslant \varliminf_{n\to\infty} \varliminf_{r\to\infty} \mathbf{P}\left\{\rho\left(g(\theta_k, \omega), g(\gamma_r, \omega)\right) > \frac{1}{n}\right\} = 0.$$

Soit $N' = \bigcup_S \bigcup_{\theta_k \in S} N(S, k)$. Alors $\mathbf{P}(N') = 0$. Si $\omega \notin N \cup N'$ et $g(\gamma, u) \in F$ pour tous les $\gamma \in \Lambda \cap G$, où G est un ouvert, et $F \subset X$ un fermé, alors pour tout $\theta_k \in G$ et toute boule S telle que $\theta_k \in S \subset G$, on a

$$g(\theta_k, \omega) \in B(S, k) \subset F.$$

D'où, par définition de l'ensemble $\{\theta_k\}$, il suit que $g(\theta, \omega) \in F$ pour tous les $\theta \in G$ et $\omega \notin N \cup N'$. Donc l'ensemble Λ réunit les conditions d'un ensemble de séparabilité d'une fonction aléatoire. ■

§ 3. Fonctions aléatoires mesurables

On conservera les notations précédentes, c'est-à-dire Θ et X sont des espaces métriques munis respectivement des distances $r(\theta_1, \theta_2)$, $\rho(x_1, x_2)$, $g(\theta, \omega)$ une fonction aléatoire définie sur Θ et à valeurs dans X, ω un événement élémentaire de l'espace probabilisé $\{\Omega, \mathfrak{S}, \mathbf{P}\}$.

Soient $\mathfrak{K}$ la tribu des boréliens de Θ, m une mesure complète sur $\mathfrak{K}$. Par $\sigma\{\mathfrak{K} \times \mathfrak{S}\}$ on désigne la plus petite tribu engendrée dans $\Theta \times \Omega$ par le produit des tribus $\mathfrak{K}$ et $\mathfrak{S}$ et par $\widetilde{\sigma}\{\mathfrak{K} \times \mathfrak{S}\}$ son complémentaire relativement à la mesure $m \times \mathbf{P}$ (cf. chapitre II, § 2).

DÉFINITION. *On dit qu'une fonction aléatoire* $g(\theta, \omega)$ *est mesurable si elle l'est par rapport à* $\widetilde{\sigma}\{\mathfrak{K} \times \mathfrak{S}\}$.

Par définition la fonction aléatoire $g(\theta, \omega)$ est $\mathfrak{S}$-mesurable $\forall\theta \in \Theta$. Le théorème de Fubini nous apprend qu'elle est $\mathfrak{K}$-mesurable comme fonction de θ $\mathbf{P}$-presque pour tous les ω. Autrement dit, ses réalisations sont presque sûrement $\mathfrak{K}$-mesurables.

Etudions maintenant les conditions d'existence d'une fonction mesurable et séparable, stochastiquement équivalente à une fonction donnée.

THÉORÈME 1. *Soient* Θ *et* X *des compacts et une mesure* m *finie. Si pour* m*-presque tous les* θ *la fonction aléatoire* $g(\theta, \omega)$ *est stochastiquement continue, il existe une fonction* $g^*(\theta, \omega)$ *aléatoire séparable mesurable, stochastiquement équivalente à* $g(\theta, \omega)$.

En vertu du théorème 1, § 2, il existe une fonction aléatoire $\tilde{g}(\theta, \omega)$ séparable, stochastiquement équivalente à une fonction $g(\theta, \omega)$. Soit I ensemble de séparabilité de la fonction $\tilde{g}(\theta, \omega)$.

Comme précédemment (§ 2), $\widetilde{A}(G, \omega)$ représente la fermeture de l'ensemble des valeurs de $\widetilde{g}(\theta, \omega)$ lorsque θ parcourt l'ensemble $G \cap I$ et $\widetilde{A}(\theta, \omega)$ l'intersection de tous les ensembles de la forme $A(S, \omega)$, où S est une boule ouverte de V (§ 2) contenant θ. En vertu de la séparabilité, $\widetilde{g}(\theta, \omega) \in A(\theta, \omega)$ presque sûrement (c'est-à-dire pour $\omega \notin N$, où $\mathsf{P}(N) = 0$). D'autre part, si la fonction $g'(\theta, \omega)$ est telle que $g'(\theta, \omega) = \widetilde{g}(\theta, \omega)$ pour $\theta \in I$ et $g'(\theta, \omega) \in \widetilde{A}(\theta, \omega)$ ($\omega \notin N'$, $\mathsf{P}(N') = 0$), alors $g'(\theta, \omega)$ est aussi une fonction aléatoire séparable (lemme 1, § 2). Construisons une fonction $g^*(\theta, \omega)$ possédant la propriété indiquée, stochastiquement équivalente à $\widetilde{g}(\theta, \omega)$ et $\widetilde{\sigma}\{\mathfrak{R} \times \mathfrak{S}\}$-mesurable. Formons la suite $\{\theta_1, \theta_2, \ldots, \theta_n, \ldots\}$ avec les points de I et soit $r_n = \min\{r(x_k, x_s), k, s, = 1, 2, \ldots, n\}$. Pour tout n construisons un recouvrement fini de l'ensemble Θ par les boules $S_1^n, \ldots, S_{j_n}^n$ de rayons r_n/n centrées en les points θ_j^n. On suppose que $\theta_j^n = \theta_j$ pour $j = 1, \ldots, n$, les autres points θ_j^n ($j = n + 1, \ldots, j_n$) sont arbitrairement choisis dans I pourvu que les boules correspondantes forment un recouvrement de Θ.

Posons $\bar{g}_n(\theta, \omega) = \widetilde{g}(\theta_j, \omega)$ pour $j \in S_k^n$, $k = 1, 2, \ldots, n$ (ces boules étant disjointes, la définition est correcte) et

$$\bar{g}_n(\theta, \omega) = \widetilde{g}(\theta_j^n, \omega) \text{ si } \theta \in S_j^n \setminus \bigcup_{i=1}^{j-1} S_i^n, \; j = n+1, \ldots, r_n.$$

On remarquera que $\bar{g}_{n+p}(\theta_n, \omega) = \widetilde{g}(\theta_n, \omega)$, $\bar{g}_n(\theta, \omega)$ est une fonction borélienne de θ à ω fixe et une fonction $\sigma\{\mathfrak{R} \times \mathfrak{S}\}$-mesurable de (θ, ω). Par ailleurs

$$\rho[\bar{g}_n(\theta, \omega), \widetilde{g}(\theta, \omega)] = \rho[\widetilde{g}(\theta_k^n, \omega), \widetilde{g}(\theta, \omega)]$$

et

$$r(\theta_k^n, \theta) < \frac{r_n}{n}. \tag{1}$$

Si l'on pose

$$G_{np}(\theta) = \mathsf{P}\{\omega : \rho[\bar{g}_n(\theta, \omega), \bar{g}_{n+p}(\theta, \omega)] > \varepsilon\},$$

d'après les conditions du théorème, $G_{np}(\theta) \to 0$ lorsque $n \to \infty$, $\forall \varepsilon > 0$, m-presque pour tous les θ. Donc

$$(m \times \mathsf{P})\{(\theta, \omega) : \rho[\bar{g}_n(\theta, \omega), \bar{g}_{n+p}(\theta, \omega)] > \varepsilon\} = \\ = \int_{\Theta} G_{np}(\theta)\, m(d\theta) \to 0 \quad \text{lorsque } n \to \infty$$

pour tous les $\varepsilon > 0$, c'est-à-dire la suite $\bar{g}_n(\theta, \omega)$ est fondamentale pour la mesure $m \times \mathsf{P}$. On peut en extraire une suite partielle

$g_{n_k}(\theta, \omega)$ convergente $m \times \mathsf{P}$-presque partout vers une fonction $\bar{g}(\theta, \omega)$ $\sigma\{\mathfrak{K} \times \mathfrak{S}\}$-mesurable. Appelons K l'ensemble des points $\{\theta, \omega\}$, où cette convergence n'a pas lieu. On peut admettre par construction que $(\theta_j, \omega) \notin K$. De toute évidence, si $(\theta, \omega) \notin K$, alors $\bar{g}(\theta, \omega) \in \tilde{A}(\theta, \omega)$. Si L est l'ensemble de points en lesquels la fonction $g(\theta, \omega)$ n'est pas stochastiquement continue, il suit de (1) que pour $\theta \notin L$

$$\mathsf{P}\{\bar{g}(\theta, \omega) \neq \tilde{g}(\theta, \omega)\} = 0. \qquad (2)$$

Posons maintenant

$$g^*(\theta, \omega) = \tilde{g}(\theta, \omega) \text{ si } \theta \in L \text{ ou si } (\theta, \omega) \in K;$$

$$g^*(\theta, \omega) = \bar{g}(\theta, \omega) \text{ si } \theta \notin L \text{ et } (\theta, \omega) \notin K.$$

Etant confondue avec $\bar{g}(\theta, \omega)$ $m \times \mathsf{P}$-presque pour tous les (θ, ω) la fonction $g^*(\theta, \omega)$ est $\tilde{\sigma}\{\mathfrak{K} \times \mathfrak{S}\}$-mesurable. Pour $\theta \in I$, $g^*(\theta, \omega) = \tilde{g}(\theta, \omega)$, de sorte que $A^*(\theta, \omega) = \tilde{A}(\theta, \omega)$, où $A^*(\theta, \omega)$ est l'ensemble $A(\theta, \omega)$ défini auparavant et construit à l'aide de la fonction $g^*(\theta, \omega)$ et de l'ensemble I, $g^*(\theta, \omega) \in A(\theta, \omega) = A^*(\theta, \omega)$. La fonction $g^*(\theta, \omega)$ est donc séparable. Eu égard à (2) on obtient

$$\mathsf{P}\{g^*(\theta, \omega) \neq \tilde{g}(\theta, \omega)\} = 0 \quad \text{pour tous les } \theta \in \Theta,$$

c'est-à-dire $g^*(\theta, \omega)$ et $g(\theta, \omega)$ sont stochastiquement équivalentes. ■

Faisons quelques remarques généralisant le théorème 1.

Remarque 1. Dans le théorème 1 la compacité des espaces Θ et X peut être remplacée par la compacité locale et la séparabilité. En effet, la compacité de l'espace X n'a servi qu'à renvoyer au théorème 1, § 2. On peut désormais se référer au théorème 2, § 2. Ceci étant, la représentation séparable et mesurable $g^*(\theta, \omega)$ de la fonction $g(\theta, \omega)$ prend généralement ses valeurs dans une extension topologique compacte de l'espace X. Si l'espace Θ est localement compact et séparable, on peut le représenter par la somme d'un nombre dénombrable de compacts. Le théorème étant valable pour chaque terme de la somme, il l'est pour leur union. Bien plus, la mesure m n'est pas forcée d'être finie, il suffit qu'elle soit σ-finie.

Remarque 2. Le théorème 1 est valable dans le cas où Θ et X sont des espaces euclidiens finidimensionnels et la mesure $\{m, \mathfrak{K}\}$, la mesure de Lebesgue sur Θ.

La démonstration du théorème 1 se simplifierait sans la séparabilité de la représentation mesurable de la fonction aléatoire. On n'aurait pas alors à faire appel à l'ensemble I. De toutes les propriétés de l'espace X seule la complétude serait utilisée.

REMARQUE 3. Si X est un espace mesurable complet, Θ un espace séparable localement compact, m une mesure σ-finie sur la tribu des boréliens de Θ, la fonction aléatoire $g(\theta, \omega)$ à valeurs dans X, $\theta \in \Theta$, $\omega \in \Omega$, stochastiquement continue m-presque pour tous les θ, est stochastiquement équivalente à une fonction aléatoire mesurable.

Le résultat important suivant découle immédiatement du théorème de Fubini.

THÉORÈME 2. *Soit* $\xi(\theta) = g(\theta, \omega)$ *une fonction aléatoire mesurable à valeurs réelles. Si*

$$\int_{\Theta} \mathsf{E}|\xi(\theta)|\, m(d\theta) < \infty,$$

alors $\forall B \in \mathfrak{R}$

$$\int_B \mathsf{E}\xi(\theta)\, m(d\theta) = \mathsf{E} \int_B \xi(\theta)\, m(d\theta).$$

La dernière égalité traduit la permutabilité de l'espérance mathématique et de l'intégration par rapport au paramètre.

§ 4. Critères de non-existence de discontinuités de seconde espèce

Fonctions sans discontinuités de seconde espèce. Soit $\xi(t)$, $t \in [a, b]$ un processus aléatoire à valeurs dans un espace métrique complet X.

DÉFINITION. *Si les réalisations du processus possèdent presque sûrement des limites à gauche et à droite pour tout* $t \in \,]a, b[$ *et au point* a (*resp.* b) *une limite à droite* (*resp. à gauche*), *on dit que le processus ne présente pas de discontinuités de seconde espèce sur l'intervalle* $[a, b]$.

Dans ce paragraphe on admettra que le processus $\xi(t)$ est séparable. Désignons par I l'ensemble de séparabilité.

DÉFINITION. *Une fonction* $x = f(t)$, $x \in X$, *possède au moins* m ε-*oscillations* ($\varepsilon > 0$) *sur* $[a, b]$ *si existent des points* $t_0, \ldots, t_m$, $a \leqslant t_0 < t_1 < \ldots < t_m \leqslant b$, *tels que* $\rho(f(t_{k-1}), f(t_k)) > \varepsilon$, $k = 1, 2, \ldots, m$.

LEMME 1. *Pour qu'une fonction* $y = f(t)$ *ne présente pas de discontinuités de seconde espèce sur* $[a, b]$ *il faut et il suffit que pour tout* $\varepsilon > 0$ *elle y possède un nombre fini d'*ε-*oscillations.*

D é m o n s t r a t i o n. *Condition suffisante.* Prouvons l'existence de la limite $f(t - 0)$ pour tout $t \in \,]a, b]$. Soit $\{t_n\}$ une suite quelconque $t_n \uparrow t$. Il ne peut exister qu'un ensemble fini de nombres t_{n_k} $(n_k < n_{k+1})$ tels que $\rho(f(t_{n_k}), f(t_{n_{k+1}})) > \varepsilon$. Par suite, à

partir d'un m on a $\rho(f(t_n), f(t_{n+k})) \leqslant 2\varepsilon$ pour tous les $n \geqslant m$, $k > 0$, c'est-à-dire la suite $f(t_n)$ est convergente. D'où l'existence de $f(t-0) = \lim\limits_{s \uparrow t} f(s)$. On démontre de façon analogue l'existence de $f(t+0)$ sur $[a, b[$.

Condition nécessaire. Supposons qu'en un point t_0 n'existe qu'une limite à gauche. On peut alors exhiber une suite $t_n \uparrow t_0$ telle que pour tout n on ait $\sup\limits_{m>n} \rho(f(t_m), f(t_n)) > \varepsilon$, c'est-à-dire le nombre d'ε-oscillations est illimité. ■

On notera que la définition du nombre d'ε-oscillations se généralise trivialement aux fonctions aléatoires étudiées sur un ensemble quelconque de valeurs réelles de t.

Dans la suite, lorsqu'on envisagera des fonctions sans discontinuités de seconde espèce, on ne fera pas de distinction entre deux fonctions possédant en chaque point $t \in [a, b]$ des limites égales à droite et à gauche. Il semble donc naturel de convenir une fois pour toute de la valeur prise par ces fonctions en un point de discontinuité. Notons $\mathscr{D}[a, b] = \mathscr{D}[a, b; X]$ l'espace des fonctions définies sur $[a, b]$ à valeurs dans X ne présentant pas de discontinuités de seconde espèce et unilatéralement continues en chaque point $t \in [a, b]$. Posons

$$\begin{aligned} \Delta_c(f) = \sup \{\min [\rho(f(t'), f(t)), \rho(f(t''), f(t))]; \\ t - c \leqslant t' < t < t'' \leqslant t + c,\ t', t, t'' \in [a, b]\} + \\ + \sup \{\rho(f(t), f(a)); a < t < a + c\} + \\ + \sup \{\rho(f(t), f(b)); b - c < t < b\}. \qquad (1) \end{aligned}$$

LEMME 2. *Pour qu'une fonction $x = f(t)$ ne présente pas de discontinuités de seconde espèce il est nécessaire et suffisant que*

$$\lim_{c \to 0} \Delta_c(f) = 0. \qquad (2)$$

Démonstration. *Condition nécessaire.* De la définition il suit que les deux derniers termes du second membre de (1) tendent vers 0 avec c pour chaque fonction $f \in \mathscr{D}[a, b]$.

Supposons que (2) n'est pas réalisée. Il existe alors des suites t'_n, t_n, t''_n telles que $t'_n < t_n < t''_n$, $t''_n - t'_n \to 0$ et $\rho(f(t'_n), f(t_n)) > \varepsilon$, $\rho(f(t''_n), f(t_n)) > \varepsilon$ pour un $\varepsilon > 0$. On peut admettre que t_n converge vers un t_0 (le cas échéant, on remplace t_n par suite partielle convergente). Des trois suites $\{t'_n\}$, $\{t_n\}$, $\{t''_n\}$ deux au moins possèdent une infinité de points situés d'un côté de t_0. Si par exemple $\{t'_n\}$ et $\{t_n\}$ sont situées à gauche de t_0, alors $f(t_n) \to f(t-0)$, $f(t'_n) \to f(t-0)$, ce qui est contraire à l'hypothèse $\rho(f(t'_n), f(t_n)) > \varepsilon$. Il en va de même lorsque $\{t_n\}$ et $\{t''_n\}$ possèdent une infinité de valeurs situées à droite de t_0. Les autres cas se ramènent aux deux premiers.

Condition suffisante. La condition (2) entraîne que $f(t)$ est continue à droite au point a et à gauche au point b. Si pour un $t_0 \in]a, b[$ n'existait pas $f(t_0+0)$, on exhiberait une suite $t_n \downarrow t$ et un $\varepsilon > 0$ tels que $\rho(f(t_n), f(t_{n+1})) > \varepsilon$, ce qui contredit (2). Donc $f(t_0+0)$ existe pour tout $t_0 \in [a, b[$. L'existence de $f(t_0-0)$ se démontre de façon analogue. La relation (2) implique que soit $f(t_0) = f(t_0-0)$, soit $f(t_0) = f(t_0+0)$. D'où le lemme. ■

Quelques inégalités.

LEMME 3. *Soient* $\xi(t)$, $t \in [0, T]$, *un processus stochastiquement continu séparable à valeurs dans* X, $g(h)$ *une fonction monotone croissante non négative et* $q(C, h) \geqslant 0$, $h \geqslant 0$, *une fonction telles que*

$$\mathsf{P}\{[\rho(\xi(t), \xi(t-h)) > Cg(h)] \cap \cap [\rho(\xi(t+h), \xi(t)) > Cg(h)]\} \leqslant q(C, h), \quad (3)$$

et

$$G = \sum_{n=0}^{\infty} g(T2^{-n}) < \infty, \quad Q(C) = \sum_{n=1}^{\infty} 2^n q(C, T2^{-n}) < \infty. \quad (4)$$

Alors pour tous les $N > 0$

$$\mathsf{P}\{\sup_{t', t'' \in [0, T]} \rho(\xi(t'), \xi(t'')) > N\} \leqslant \leqslant \mathsf{P}\left\{\rho(\xi(0), \xi(T)) > \frac{N}{2G}\right\} + Q\left(\frac{N}{2G}\right).$$

D é m o n s t r a t i o n. Posons

$$A_{nk} = \left\{\rho\left(\xi\left(\frac{k+1}{2^n}T\right), \xi\left(\frac{k}{2^n}T\right)\right) \leqslant Cg(T2^{-n})\right\},$$
$$k = 0, 1, \ldots, 2^n - 1, \; n = 0, 1, 2, \ldots,$$

$$B_{nk} = A_{nk-1} \cup A_{nk}, \quad D_n = \bigcap_{m=n}^{\infty} \bigcap_{k=1}^{2^m-1} B_{mk} \; (n \geqslant 1),$$

$$D_0 = A_{00} \cap D_1.$$

La continuité stochastique entraîne que (théorème 3, § 2) l'ensemble de séparabilité I du processus $\xi(t)$ est un ensemble de nombres de la forme $k/2^n$, $k = 0, 1, 2, \ldots, n = 0, 1, 2, \ldots$ On a

$$\mathsf{P}\{\bar{D}_n\} \leqslant \sum_{m=n}^{\infty} \sum_{k=1}^{2^m-1} P\{\bar{B}_{mk}\} \leqslant \sum_{m=n}^{\infty} 2^m q(C, T2^{-m}) = Q(n, C), \quad (5)$$

où

$$Q(n, C) = \sum_{m=n}^{\infty} 2^m q(C, T2^{-m}).$$

De la relation $D_0 = A_{00} \cap D_1$ il suit que $\rho(\xi(T), \xi(0)) \leqslant Cg(T)$ et l'un des deux événements a lieu: ou bien $\rho(\xi(T/2), \xi(0)) \leqslant Cg(T2^{-1})$, ou bien $\rho(\xi(T), \xi(T2^{-1})) \leqslant Cg(T2^{-1})$. Dans les deux cas

$$\rho(\xi(0), \xi(T/2)) \leqslant Cg(T) + Cg(T2^{-1}),$$
$$\rho(\xi(T/2), \xi(T)) \leqslant Cg(T) + Cg(T2^{-1}).$$

Raisonnons par récurrence. Supposons que

$$\rho\left(\xi\left(\frac{k}{2^m} T\right), \xi\left(\frac{j}{2^m} T\right)\right) < Cg(T) + 2C \sum_{s=1}^{m} g(T2^{-s}) \qquad (6)$$

est vraie pour $m = n$ et pour $k, j = 0, 1, \ldots, 2^n$ sous l'hypothèse que l'événement D_0 a lieu. Montrons que cette inégalité est vraie pour $m = n + 1$. Soient k et j deux nombres impairs: $k = 2k_1 + 1$, $j = 2j_1 + 1$. Etant donné que de D_{n+1} il suit qu'au moins une des inégalités

$$\rho\left(\xi\left(\frac{k_1}{2^n} T\right), \xi\left(\frac{2k_1+1}{2^{n+1}} T\right)\right) \leqslant Cg(T2^{-(n+1)}),$$
$$\rho\left(\xi\left(\frac{k_1+1}{2^n} T\right), \xi\left(\frac{2k_1+1}{2^{n+1}} T\right)\right) \leqslant Cg(T2^{-(n+1)})$$

est réalisée, on a

$$\rho\left(\xi\left(\frac{k}{2^{n+1}} T\right), \xi\left(\frac{k'}{2^n} T\right)\right) \leqslant Cg(T2^{-(n+1)}),$$

où k' est égal soit à k_1, soit à $k_1 + 1$. De façon analogue il existe un j' tel que

$$\rho\left(\xi\left(\frac{j}{2^{n+1}} T\right), \xi\left(\frac{j'}{2n} T\right)\right) \leqslant Cg(T2^{-(n+1)}).$$

Le raisonnement par récurrence donne

$$\rho\left(\xi\left(\frac{k}{2^{n+1}} T\right), \xi\left(\frac{j}{2^{n+1}} T\right)\right) \leqslant Cg(T) + 2C \sum_{s=1}^{n+1} g(T2^{-s}).$$

On traite de façon analogue le cas où k ou j sont pairs. Donc l'inégalité (5) est vraie pour tous les $m \geqslant 1$. La séparabilité du processus implique que si l'événement D_0 est réalisé, alors

$$\sup \{\rho(\xi(t'), \xi(t'')),\ t', t'' \in [0, T]\} \leqslant 2CG$$

presque sûrement. D'où il suit

$$\mathsf{P}\{\sup_{t', t'' \in [0, T]} \rho(\xi(t'), \xi(t'')) > N\} \leqslant$$
$$\leqslant Q\left(\frac{N}{2G}\right) + \mathsf{P}\left\{\rho(\xi(0), \xi(T)) > \frac{N}{2G}\right\}. \quad ■$$

LEMME 4. *Sous les conditions du lemme précédent on a*

$$\mathsf{P}\left\{\Delta_\varepsilon(\xi) > CG\left(\left[\log_2 \frac{T}{2\varepsilon}\right]\right)\right\} \leqslant Q\left(\left[\log_2 \frac{T}{2\varepsilon}\right], C\right), \tag{7}$$

où

$$G(n) = \sum_{m=n}^{\infty} g(T2^{-m}), \quad Q(n, C) = \sum_{m=n}^{\infty} 2^m q(C, 2T^{-m}).$$

Démonstration. Poursuivons les raisonnements du lemme précédent. Supposons que l'événement D_n est réalisé. En raisonnant par récurrence on démontre que pour tout k et m il existe un j_{nm} $(0 \leqslant j_{nm} < 2^{m+1})$ entier tel que

$$\max_{0 \leqslant j \leqslant j_{nm}} \rho\left(\xi\left(\frac{k-1}{2^n} T\right), \xi\left(\left[\frac{k-1}{2^n} + \frac{j}{2^{n+m}}\right] T\right)\right) \leqslant C \sum_{s=n}^{n+m} g(T2^{-s}), \tag{8}$$

$$\max_{j_{nm}+1 \leqslant j \leqslant 2^{m+1}} \rho\left(\xi\left(\left[\frac{k-1}{2^n} + \frac{j}{2^{n+m}}\right] T\right), \xi\left(\frac{k+1}{2^n} T\right)\right) \leqslant$$

$$\leqslant C \sum_{s=n}^{n+m} g(T2^{-s}), \tag{9}$$

et de plus $j_{nm} 2^{-(n+m)}$ est monotone non décroissante comme fonction de m (à n et k fixes). Pour $m = 0$ il convient de prendre

$$j_{n0} = 0 \text{ si } \rho(\xi(kT2^{-n}), \xi((k+1)T2^{-n})) \leqslant Cg(T2^{-n}),$$

$$j_{n0} = 1 \text{ si } \rho(\xi((k-1)T2^{-n}), \xi(kT2^{-n})) \leqslant Cg(T2^{-n}).$$

Si l'on admet que l'événement D_n est réalisé, l'une de ces deux inégalités a forcément lieu. Supposons que j_{nm} ait déjà été choisi. Alors on prendra

$j_{nm+1} = 2j_{nm}$ si

$$\rho\left(\xi\left(\left[\frac{k-1}{2^n} + \frac{2j_{nm}+1}{2^{n+m+1}}\right] T\right), \xi\left(\left[\frac{k-1}{2^n} + \frac{j_{nm}+1}{2^{n+m}}\right] T\right)\right) \leqslant$$

$$\leqslant Cg(T2^{-(n+m+1)}),$$

$j_{nm+1} = 2j_{nm} + 1$ si

$$\rho\left(\xi\left(\left[\frac{k-1}{2^n} + \frac{j_{nm}}{2^{n+m}}\right] T\right), \xi\left(\left[\frac{k-1}{2^n} + \frac{2j_{nm}+1}{2^{n+m+1}}\right] T\right)\right) \leqslant$$

$$\leqslant Cg(T2^{-(n+m+1)}).$$

Ceci est possible puisque l'une des deux inégalités a forcément lieu si D_n est réalisé; si les deux inégalités sont réalisées, on prendra l'une des valeurs indiquées de $j_{nm} + 1$.

En passant à la limite pour $m \to \infty$ dans les relations (8) et (9), on constate que pour toute réalisation telle que D_n ait lieu il

existe $\tau = \tau(\omega)$, $0 \leqslant \tau \leqslant T2^{-(n-1)}$, tel que

$$\sup_{\substack{0<t<\tau \\ t\in I}} \rho\left(\xi\left(\frac{k-1}{2^n}T\right),\ \xi\left(\frac{k-1}{2^n}T+t\right)\right) \leqslant CG(n)$$

et

$$\sup_{\substack{\tau<t<T2^{-(n-1)} \\ t\in I}} \rho\left(\xi\left(\frac{k-1}{2^n}T+t\right),\ \xi\left(\frac{k+1}{2^n}T\right)\right) \leqslant CG(n).$$

Soit $\varepsilon \in [2^{-(n+1)}T,\ 2^{-n}T]$ et $0 < t'' - t' < \varepsilon$. Il existe k tel que

$$(k-1)2^{-n}T \leqslant t' < t'' < (k+1)\,2^{-n}T.$$

Si $t \in [t',\ t'']$, alors

$$\text{soit }]t',\ t[\ \subset [(k-1)\,2^{-n}T,\ (k-1)\,2^{-n}T+\tau],$$
$$\text{soit }]t',\ t''[\ \subset [(k-1)\,2^{-n}T+\tau,\ (k+1)\,2^{-n}T].$$

Si $t',\ t,\ t'' \in I$, l'une au moins des inégalités est vraie:

$$\rho(\xi(t'),\ \xi(t)) \leqslant 2CG(n),\quad \rho(\xi(t),\ \xi(t'')) \leqslant 2CG(n).$$

La séparabilité du processus implique que l'une de ces deux inégalités est presque sûrement vraie quelle que soit la réalisation du processus. Donc si D_n a lieu, on a presque sûrement

$$\Delta_\varepsilon(\xi) \leqslant 2CG(n).$$

En vertu de l'inégalité (5)

$$\mathsf{P}\{\Delta_\varepsilon(\xi) > 2CG(n)\} \leqslant \mathsf{P}(\bar{D}_n) \leqslant Q(n,\ C),$$

ou finalement, compte tenu de ce que $\varepsilon \geqslant 2^{-(n+1)}T$ et de la monotonie des fonctions $g(h)$ et $q(h)$,

$$\mathsf{P}\left\{\Delta_\varepsilon(\xi) > CG\left(\left[\log_2\frac{T}{2\varepsilon}\right]\right)\right\} \leqslant Q\left(\left[\log_2\frac{T}{2\varepsilon}\right],\ C\right).\quad ■$$

Conditions de non-existence de discontinuités de seconde espèce faisant intervenir les répartitions partielles du processus. Du dernier lemme on déduit immédiatement le

THÉORÈME 1. *Si $\xi(t)$, $t \in [0,\ T]$, est un processus stochastiquement continu séparable à valeurs dans X et tel que*

$$\mathsf{P}\{[\rho(\xi(t),\ \xi(t-h)) \geqslant Cg(h)] \cap [\rho(\xi(t+h),\ \xi(t)) \geqslant Cg(h)]\} \leqslant q(C,\ h), \tag{10}$$

où

$$\sum_{n=1}^{\infty} g(T2^{-n}) < \infty,\quad \sum_{n=1}^{\infty} 2^n q(C,\ T2^{-n}) < \infty, \tag{11}$$

alors $\xi(t)$ ne présente presque sûrement pas de discontinuités de seconde espèce.

D é m o n s t r a t i o n. En posant $C = 1$ dans l'inégalité (7) on voit que, dans les conditions du théorème, $\Delta_\varepsilon(\xi) \to 0$ stochastiquement pour $\varepsilon \to 0$. Or $\Delta_\varepsilon(\xi)$ est monotone décroissante comme fonction de ε pour $\varepsilon \downarrow 0$. Donc $\lim \Delta_\varepsilon(\xi)$ pour $\varepsilon \to 0$ existe presque sûrement et vaut zéro. ■

COROLLAIRE. *Soit donné sur* $[0, T]$ *un processus aléatoire stochastiquement continu au sens large, à valeurs dans un espace X complet séparable localement compact, dont les répartitions partielles* (« *tridimensionnelles* ») *vérifient les conditions* (10), (11). *Il existe alors une représentation de ce processus ne présentant pas de discontinuités de seconde espèce.*

Conditions de non-existence de discontinuités de seconde espèce faisant intervenir les probabilités conditionnelles. Dans le théorème précédent la condition de non-existence de discontinuités de seconde espèce s'exprimait en fonction de propriétés des répartitions partielles (« tridimensionnelles ») du processus aléatoire. On se propose maintenant d'exhiber des résultats d'une autre nature, utilisant des hypothèses relatives aux probabilités conditionnelles et applicables lorsqu'on dispose d'une information importante sur les propriétés des répartitions conditionnelles du processus.

Soit $\{\mathfrak{F}_t, t \in [0, T]\}$ un flot de tribus. Supposons que le processus $\xi(t)$ est adapté au flot $\{\mathfrak{F}_t, t \in [0, T]\}$, c'est-à-dire l'élément aléatoire $\xi(t)$ est $\mathfrak{F}_t$-mesurable pour tout $t \in [0, T]$.

Soit

$$\alpha(\varepsilon, \delta) = \inf \sup_{s,t} [\mathsf{P}\{\rho(\xi(s), \xi(t)) \geqslant \varepsilon \mid \mathfrak{F}_s\}; \\ 0 \leqslant s \leqslant t \leqslant s + \delta \leqslant T, \omega \in \Omega'], \qquad (12)$$

où inf est pris sur tous les sous-ensembles Ω' ($\Omega' \in \mathfrak{S}$) de probabilité 1. On remarque sans peine qu'il existe Ω^0, $\mathsf{P}(\Omega^0) = 1$, $\Omega^0 \in \mathfrak{S}$, sur lequel l'infimum envisagé est réalisé, tel que

$$\alpha(\varepsilon, \delta) = \sup_{s,t} \{\mathsf{P}\{\rho(\xi(s), \xi(t)) \geqslant \varepsilon \mid \mathfrak{F}_s\}; \\ 0 \leqslant s \leqslant t \leqslant s + \delta \leqslant T, \omega \in \Omega^0\}.$$

Montrons que pour les processus séparables la condition $\alpha(\varepsilon, \delta) \to 0$ pour $\delta \to 0$, $\forall \varepsilon > 0$, est une condition de non-existence de discontinuités de seconde espèce. Soient $[c, d]$ un intervalle fixe, $[c, d] \subset [0, T]$, I une suite finie d'instants $t_1, t_2, \ldots, t_n$, $s \leqslant c \leqslant \leqslant t_1 < t_2 < \ldots < t_n \leqslant d$. Soit $A(\varepsilon, Z)$ l'événement : la réalisation du processus $\xi(t)$ possède au moins une ε-oscillation sur $[c, d] \cap \cap Z$.

LEMME 5. *On a presque sûrement*

$$\mathsf{P}\{A(\varepsilon, I) \mid \mathfrak{F}_s\} \leqslant 2\alpha\left(\frac{\varepsilon}{4}, d - c\right). \qquad (13)$$

D é m o n s t r a t i o n. On remarquera tout d'abord que les propriétés des espérances mathématiques conditionnelles entraînent

pour $s < t < u$

$$\mathsf{P}\{\rho(\xi(t), \xi(u)) \geqslant \varepsilon \mid \mathfrak{F}_s\} = \mathsf{E}\{\mathsf{P}\{\rho(\xi(t), \xi(u)) \geqslant \varepsilon \mid \mathfrak{F}_t\} \mid \mathfrak{F}_s\} \leqslant \alpha(\varepsilon, u-s). \tag{14}$$

Soient les événements

$$B_k = \left\{\rho(\xi(c), \xi(t_i)) < \frac{\varepsilon}{2}, \; i=1, 2, \ldots, k-1, \; \rho(\xi(c), \xi(t_k)) > \frac{\varepsilon}{2}\right\},$$

$$C_k = \left\{\rho(\xi(t_k), \xi(d)) \geqslant \frac{\varepsilon}{4}\right\}, \quad D_k = B_k \cap C_k, \quad k = 1, 2, \ldots, n,$$

$$C_0 = \left\{\rho(\xi(c), \xi(d)) \geqslant \frac{\varepsilon}{4}\right\}.$$

Les événements B_k sont incompatibles et si l'on pose $D = \bigcup_{k=1}^{n} D_k$, alors $A(\varepsilon, I) \subset C_0 \cup D$. En effet, si $A(\varepsilon, I)$ a lieu, l'inégalité

$$\rho(\xi(c), \xi(t_k)) \geqslant \frac{\varepsilon}{2}$$

est réalisée pour la première fois pour un k, c'est-à-dire est réalisé l'un des événements B_k $(k = 1, \ldots, n)$. Si de plus D n'a pas lieu, c'est-à-dire si

$$\rho(\xi(t_k), \xi(d)) < \frac{\varepsilon}{4},$$

alors

$$\rho(\xi(c), \xi(d)) \geqslant \rho(\xi(c), \xi(t_k)) - \rho(\xi(t_k), \xi(d)) > \frac{\varepsilon}{4},$$

c'est-à-dire l'événement C_0 a lieu. Donc $A(\varepsilon, I) \subset C_0 \cup D$. On a presque sûrement

$$\mathsf{P}\{D_k \mid \mathfrak{F}_s\} = \mathsf{E}\{\chi_{D_k} \mid \mathfrak{F}_s\} = \mathsf{E}\{\mathsf{E}\{\chi_{B_k}\chi_{C_k} \mid \mathfrak{F}_{t_k}\} \mid \mathfrak{F}_s\} =$$
$$= \mathsf{E}\{\chi_{B_k}\mathsf{P}\{C_k \mid \mathfrak{F}_{t_k}\} \mid \mathfrak{F}_s\} \leqslant \alpha\left(\frac{\varepsilon}{4}, d-c\right) \mathsf{E}\{\chi_{B_k} \mid \mathfrak{F}_s\},$$

où χ_A désigne toujours l'indicateur de l'événement A. D'où il suit

$$\mathsf{P}\{D \mid \mathfrak{F}_s\} = \sum_{k=1}^{n} \mathsf{P}\{D_k \mid \mathfrak{F}_s\} \leqslant \alpha\left(\frac{\varepsilon}{4}, d-c\right) \mathsf{E}\left\{\sum_{k=1}^{n} \chi_{B_k} \mid \mathfrak{F}_s\right\} \leqslant$$
$$\leqslant \alpha\left(\frac{\varepsilon}{4}, d-c\right) \pmod{\mathsf{P}}.$$

En vertu de (14) $\mathsf{P}\{C_0 \mid \mathfrak{F}_s\} \leqslant \alpha\left(\frac{\varepsilon}{4}, d-c\right)$. Donc

$$\mathsf{P}\{A(\varepsilon, I) \mid \mathfrak{F}_s\} \leqslant \mathsf{P}\{D \mid \mathfrak{F}_s\} + \mathsf{P}\{C_0 \mid \mathfrak{F}_s\} \leqslant 2\alpha\left(\frac{\varepsilon}{4}, d-c\right) \pmod{\mathsf{P}},$$

ce qui achève la démonstration du lemme.

LEMME 6. *Soit $A^k(\varepsilon, I)$ l'événement: $\xi(t)$ possède au moins k ε-oscillations sur I. Alors*

$$\mathsf{P}\{A^k(\varepsilon, I)|\mathfrak{F}_s\} \leqslant \left[2\alpha\left(\frac{\varepsilon}{4}, d-c\right)\right]^k \pmod{\mathsf{P}}. \tag{15}$$

D é m o n s t r a t i o n. Soit $B_r(\varepsilon, I)$ l'événement: la réalisation du processus $\xi(t)$ possède au moins $k-1$ ε-oscillations sur l'ensemble $(t_1, \ldots, t_r)$ et moins de $k-1$ ε-oscillations sur $(t_1, \ldots, t_{r-1})$. Les événements $B_r(\varepsilon, I)$ $(r = 1, \ldots, n)$ sont incompatibles et $\bigcup_{r=1}^{n} B_r(\varepsilon, I) = A^{k-1}(\varepsilon, I) \supset A^k(\varepsilon, I)$. Par ailleurs, $A^k(\varepsilon, I) \cap B_r(\varepsilon, I)$ implique que la réalisation possède au moins une ε-oscillation sur l'ensemble $(t_r, t_{r+1}, \ldots, t_n)$. Par suite

$$A^k(\varepsilon, I) \subset \bigcup_{r=1}^{n} (B_r(\varepsilon, I) \cap C_r(\varepsilon, I)),$$

où $C_r(\varepsilon, I)$ signifie que $\xi(t)$ possède au moins une ε-oscillation sur $(t_r, t_{r+1}, \ldots, t_n)$. Donc

$$\mathsf{P}\{A^k(\varepsilon, I)|\mathfrak{F}_s\} \leqslant \sum_{r=1}^{n} \mathsf{P}\{B_r(\varepsilon, I) \cap C_r(\varepsilon, I)|\mathfrak{F}_s\} \pmod{\mathsf{P}}. \tag{16}$$

Les propriétés des espérances mathématiques conditionnelles donnent

$$\mathsf{P}\{B_r(\varepsilon, I) \cap C_r(\varepsilon, I)|\mathfrak{F}_s\} = \mathsf{E}\{\mathsf{E}\{\chi_{B_r(\varepsilon, I)}\chi_{C_r(\varepsilon, I)}|\mathfrak{F}_{t_r}\}|\mathfrak{F}_s\} \leqslant$$
$$\leqslant \mathsf{E}\{\chi_{B_r(\varepsilon, I)}\mathsf{P}\{C_r(\varepsilon, I)|\mathfrak{F}_{t_r}\}|\mathfrak{F}_s\} \leqslant$$
$$\leqslant 2\alpha\left(\frac{\varepsilon}{4}, d-c\right)\mathsf{P}\{B_r(\varepsilon, I)|\mathfrak{F}_s\} \pmod{\mathsf{P}}.$$

Ceci et (16) entraînent

$$\mathsf{P}\{A^k(\varepsilon, I)|\mathfrak{F}_s\} \leqslant 2\alpha\left(\frac{\varepsilon}{4}, d-c\right)\sum_{r=1}^{n} \mathsf{P}\{B_r(\varepsilon, I)|\mathfrak{F}_s\} =$$
$$= 2\alpha\left(\frac{\varepsilon}{4}, d-c\right)\mathsf{P}\{A^{k-1}(\varepsilon, I)|\mathfrak{F}_s\} \pmod{\mathsf{P}},$$

d'où suit la proposition annoncée. ■

THÉORÈME 2. *Si $\xi(t)$ est un processus séparable et pour tout $\varepsilon > 0$*

$$\lim_{\delta\to 0} \alpha(\varepsilon, \delta) = 0, \tag{17}$$

alors il ne présente pas de discontinuités de seconde espèce.

Il suffit de prouver que toute réalisation de $\xi(t)$ possède presque sûrement un nombre fini d'ε-oscillations. Soit I ensemble de séparabilité du processus $\xi(t)$. Ecrivons cet ensemble sous la forme $I = \bigcup_{n=1}^{\infty} I_n$, où I_n est une suite monotone croissante d'ensembles composés d'un nombre fini d'éléments. Soit donné $\varepsilon > 0$. Parta-

geons $[0, T]$ en m intervalles Δ_r, $r = 1, \ldots, m$, de même longueur de telle sorte que $2\alpha\left(\frac{\varepsilon}{4}, \frac{T}{m}\right) = \beta < 1$. Alors

$$\mathsf{P}\{A^{\infty}(\varepsilon, I \cap \Delta_r) | \mathfrak{F}_s) \leqslant \mathsf{P}\{A^k(\varepsilon, J \cap \Delta_r) | \mathfrak{F}_s\} = \\ = \lim_{n\to\infty} \mathsf{P}\{A^k(\varepsilon, I_n \cap \Delta_r) | \mathfrak{F}_s\} \leqslant \beta^k,$$

d'où

$$\mathsf{P}\{A^{\infty}(\varepsilon, I \cap \Delta_r) | \mathfrak{F}_s\} = 0 \pmod{\mathsf{P}} \text{ et } \mathsf{P}\{A^{\infty}(\varepsilon, I \cap \Delta_r)\} = 0.$$

Donc

$$\mathsf{P}\{A^{\infty}(\varepsilon, I)\} = 0. \quad \blacksquare$$

Signalons quelques conséquences importantes du théorème précédent.

THÉORÈME 3. *Tout processus* $\xi(t)$, $t \in [0, T]$ *stochastiquement continu séparable à accroissements indépendants et à valeurs dans un espace vectoriel normé* X *ne présente pas de discontinuités de seconde espèce.*

En effet, de la définition des processus à accroissements indépendants il suit que

$$\mathsf{P}\{|\xi(s) - \xi(t)| \geqslant \varepsilon | \mathfrak{F}_s\} = \mathsf{P}\{|\xi(s) - \xi(t)| \geqslant \varepsilon\} \pmod{\mathsf{P}}.$$

D'autre part, la continuité stochastique uniforme (cf. théorème 2, § 1, chapitre I) implique que

$$\alpha(\varepsilon, \delta) = \sup\{\mathsf{P}[|\xi(s) - \xi(t)| \geqslant \varepsilon];\ 0 \leqslant s \leqslant t \leqslant s + \delta \leqslant T\}$$

tend vers zéro avec δ quel que soit $\varepsilon > 0$. Donc les conditions du théorème 2 sont réalisées.

Le théorème 2 fournit de puissants résultats pour les processus markoviens.

THÉORÈME 4. *Si* $\xi(t)$, $t \in [0, T]$ *est un processus markovien séparable à valeurs dans un espace métrique* X *et à probabilités de passage* $\mathsf{P}(t, x, s, A)$ *vérifiant la condition*

$$\alpha(\varepsilon, \delta) = \sup[\mathsf{P}\{s, x, t, \bar{S}_\varepsilon(x)\};\ x \in X,\ 0 \leqslant s \leqslant t \leqslant s + \delta \leqslant T] \to 0$$

pour $\delta \to 0$, *où* $S_\varepsilon(x)$ *est la boule de rayon* ε *centrée en* x, $\bar{S}_\varepsilon(x)$ *son complémentaire, alors le processus* $\xi(t)$ *ne présente pas de discontinuités de seconde espèce.*

Ceci découle directement du théorème 2 et de la définition du processus markovien.

Régularisation des réalisations d'un processus ne présentant pas de discontinuités de seconde espèce. On a vu précédemment que pour étudier des fonctions sans discontinuités de seconde espèce il fallait identifier des fonctions admettant en chaque point des limites égales à gauche et à droite.

On rappelle que si un processus est séparable, les valeurs de la réalisation de $\xi(t)$ sont presque sûrement les valeurs de suites $\xi(t_i)$ pour $t_i \to t$, $t_i \in I$, ensemble de séparabilité. Si de plus le processus ne présente pas de discontinuités de seconde espèce, alors pour chaque t $\xi(t)$ est égal à $\xi(t-0)$ ou $\xi(t+0)$ presque sûrement.

THÉORÈME 5. *Si $\xi(t)$ est un processus stochastiquement continu à droite sans discontinuités de seconde espèce à valeurs dans un espace métrique X, alors il est équivalent à un processus $\xi'(t)$ dont les réalisations sont continues à droite* (mod P).

Démonstration. L'événement A: $\lim\limits_{n\to\infty}\xi\left(t+\frac{1}{n}\right)$ existe pour tout $t \in [0, T]$, a lieu presque sûrement. Posons $\xi'(t) = \lim\limits_{n\to\infty}\xi\left(t+\frac{1}{n}\right)$ dans le cas A et $\xi'(t)=\xi(t)$ dans $\bar{A}$. On a

$$\{\xi'(t)\neq\xi(t)\} = \bigcup_{m=1}^{\infty}\left\{\rho(\xi(t), \xi'(t)) > \frac{1}{m}\right\}\cap A,$$

$$\mathsf{P}\{\xi'(t)\neq\xi(t)\} = \lim_{m\to\infty}\mathsf{P}\left\{\left(\rho(\xi(t), \xi'(t)) > \frac{1}{m}\right)\cap A\right\}.$$

Par ailleurs,

$$\mathsf{P}\left\{\rho(\xi(t), \xi'(t)) > \frac{1}{m}\right\} = \\ = \mathsf{P}\left\{\bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty}\left\{\rho\left(\xi(t), \xi\left(t+\frac{1}{n}\right)\right) > \frac{1}{m}\right\}\right\} = \\ \lim_{k\to\infty}\mathsf{P}\left\{\bigcap_{n=k}^{\infty}\left\{\rho\left(\xi(t), \xi\left(t+\frac{1}{n}\right)\right) > \frac{1}{m}\right\}\right\} \leqslant \\ \leqslant \lim_{n\to\infty}\mathsf{P}\left\{\rho\left(\xi(t), \xi\left(t+\frac{1}{n}\right)\right) > \frac{1}{m}\right\}.$$

Donc $\mathsf{P}\{\xi'(t)\neq\xi(t)\} = 0$. Reste à noter que la fonction $\xi'(t)$ est continue à droite sur l'ensemble A. ∎

§ 5. Processus continus

Conditions de continuité d'un processus ne présentant pas de discontinuités de seconde espèce. On suppose toujours que X est un espace métrique complet, $\xi(t)$, $t \in [0, T]$, un processus aléatoire à valeurs dans X.

DÉFINITION. *On dit que le processus $\xi(t)$, $t \in [0, T]$, est continu si presque toutes ses réalisations sont continues sur* $[0, T]$.

On peut indiquer une condition de continuité suffisante simple pour les processus ne présentant pas de discontinuités de seconde espècs.

THÉORÈME 1. *Soit $\{t_{nk}, k = 0, 1, \ldots, m_n\}$, $n = 1, 2, \ldots$, une suite de partitions de l'intervalle* $[0, T]$, $0 = t_{n0} < t_{n1} < \ldots <$

$< t_{nm_n} = T$ *et* $\lambda_n = \max_{1 \leqslant k \leqslant mn} (t_{nk} - t_{n,k-1}) \to 0$ *pour* $n \to \infty$.
Si le processus $\xi(t)$ *est séparable, ne présente pas de discontinuités de seconde espèce et*

$$\sum_{k=1}^{mn} \mathsf{P}\{\rho[\xi(t_{nk}), \xi(t_{n,k-1})] > \varepsilon\} \to 0 \quad \textit{pour } \lambda_n \to 0, \tag{1}$$

alors il est continu.

D é m o n s t r a t i o n. Soit ν_ε $(0 \leqslant \nu_\varepsilon \leqslant \infty)$ le nombre de valeurs de t telles que $\rho[\xi(t+0), \xi(t-0)] > 2\varepsilon$, et $\nu_\varepsilon^{(n)}$ le nombre des indices k tels que $\rho[\xi(t_{nk}), \xi(t_{n,k-1})] < \varepsilon$. De toute évidence $\nu_\varepsilon \leqslant \varliminf_{n\to\infty} \nu_\varepsilon^{(n)}$. Par ailleurs

$$\mathsf{E}\nu_\varepsilon^{(n)} = \sum_{k=1}^{mn} \mathsf{P}\{\rho[\xi(t_{nk}), \xi(t_{n,k-1})] > \varepsilon\}.$$

Le lemme de Fatou nous dit que $\mathsf{E}\nu_\varepsilon \leqslant \mathsf{E} \varliminf_{n\to\infty} \nu_\varepsilon^{(n)} \leqslant \varliminf_{n\to\infty} \mathsf{E}\nu_\varepsilon^{(n)}$. Par suite $\mathsf{E}\nu_\varepsilon = 0$, c'est-à-dire $\nu_\varepsilon = 0$ presque sûrement pour tout $\varepsilon > 0$. Donc $\xi(t-0) = \xi(t+0)$ presque sûrement pour tout t. Le processus étant séparable, on a $\xi(t) = \xi(t-0) = \xi(t+0)$, i.e. le processus est continu. ∎

Appliquons le théorème 1 aux processus vérifiant les conditions du théorème 2, § 4. Supposons que $\alpha(\varepsilon, \delta)$ est défini par (12), § 4.

THÉORÈME 2. *Si le processus* $\xi(t)$ *est séparable et*

$$\lim_{\delta\to 0} \frac{\mathsf{E}\alpha(\varepsilon, \delta)}{\delta} = 0 \quad \forall \varepsilon > 0, \tag{2}$$

il est continu.

Le processus $\xi(t)$ ne possédant pas de discontinuités de seconde espèce si la condition (2) est réalisée, il suffit de vérifier (1). Eu égard à

$$\mathsf{P}\{\rho[\xi(t_{nk}), \xi(t_{n,k-1})] > \varepsilon\} \mathsf{E} \leqslant \mathsf{E}\alpha(\varepsilon, \Delta t_{nk}),$$

où $\Delta t_{nk} = t_{nk} - t_{n,k-1}$, on trouve

$$\sum_{k=1}^{mn} \mathsf{P}\{\rho[\xi(t_{nk}), \xi(t_{n,k-1})] > \varepsilon\} \leqslant (b-a) \max_{1 \leqslant k \leqslant n} \frac{\mathsf{E}\alpha(\varepsilon, \Delta t_{nk})}{\Delta t_{nk}} \to 0$$

pour $\lambda_n \to 0$. ∎

En appliquant le théorème 2 aux processus markoviens on obtient la condition suivante de continuité.

THÉORÈME 3. *Si* $\xi(t)$ *est un processus markovien séparable et*

$$\frac{1}{\delta} \mathsf{P}(s, y, t, \overline{S}_\varepsilon(y)) \to 0$$

pour $\delta \to 0$, $\forall \varepsilon > 0$ *fixe, uniformément en* y, s, t, $0 \leqslant t - s \leqslant \delta$, *il est continu.*

$\overline{S}_\varepsilon(x)$ est le complémentaire de la boule $S_\varepsilon(x)$ de rayon ε centrée en x.

Processus à accroissements indépendants. Le théorème 1 ne fournit que des conditions suffisantes de continuité d'un processus aléatoire. Il se trouve que les conditions du théorème sont aussi nécessaires pour un cas particulier de processus à accroissements indépendants.

THÉORÈME 4. *Si un processus* $\xi(t)$ *à accroissements indépendants est continu, la condition* (1) *est réalisée pour une suite* $\{t_{nk}, k = 0, \ldots, m_n\}$, $n = 1, 2, \ldots$, *de partitions de l'intervalle* $[0, T]$ *telle que* $\lambda_n \to 0$.

Démonstration. Posons $\Delta_h = \sup_{|t_1 - t_2| \leqslant h} \rho[\xi(t_1), \xi(t_2)]$. La continuité du processus $\xi(t)$ entraîne que $\Delta_h \to 0$ presque sûrement avec h. Donc $\lim_{h \to 0} \mathsf{P}\{\Delta_n > \varepsilon\} = 0$. D'autre part, si $\lambda_n < h$, alors

$$\mathsf{P}\{\Delta_h > \varepsilon\} \geqslant \mathsf{P}\{\sup \rho[\xi(t_{nk}), \xi(t_{n,k-1})] > \varepsilon\} =$$

$$= \mathsf{P}\{\rho[\xi(t_{n1}), \xi(t_{n0})] > \varepsilon\} + \mathsf{P}\{\rho[\xi(t_{n1}), \xi(t_{n0})] \leqslant \varepsilon\} \times$$

$$\times \mathsf{P}\{\rho[\xi(t_{n2}), \xi(t_{n1})] > \varepsilon\} + \ldots + \prod_{k=1}^{m_n - 1} \mathsf{P}\{\rho[\xi(t_{nk}), \xi(t_{n,k-1})] \leqslant \varepsilon\} \times$$

$$\times \mathsf{P}\{\rho[\xi(t_{nm_n}), \xi(t_{n,m_n-1})] > \varepsilon\} \geqslant$$

$$\geqslant \mathsf{P}\{\Delta_h \leqslant \varepsilon\} \sum_{k=1}^{m_n} \mathsf{P}\{\rho[\xi(t_{nk}), \xi(t_{n,k-1})] > \varepsilon\},$$

d'où

$$\sum_{k=1}^{m_n} \mathsf{P}\{\rho[\xi(t_{nk}), \xi(t_{n,k-1})] > \varepsilon\} \leqslant \frac{\mathsf{P}\{\Delta_h > \varepsilon\}}{\mathsf{P}\{\Delta_h \leqslant \varepsilon\}} \to 0$$

pour $h \to 0$, $\forall \varepsilon > 0$. ■

COROLLAIRE. *Une condition nécessaire et suffisante pour qu'un processus aléatoire* $\xi(t)$ *à accroissements indépendants soit continu est qu'il soit un processus de mouvement brownien, d'espérance mathématique* $a(t) = \mathsf{E}\xi(t)$ *continue et de matrice de variance* $B(t)$ *continue.*

Ce corollaire découle du théorème 4 et du théorème 1, § 3, chapitre I.

Condition de Kolmogorov de continuité d'un processus aléatoire. Prouvons une condition suffisante directe (c'est-à-dire n'utilisant pas la non-existence des discontinuités de seconde espèce)

de continuité d'un processus aléatoire. Elle est basée sur une variante simplifiée des lemmes 3 et 4, § 4.

LEMME 1. *Soit $\xi(t)$, $t \in [0, T]$, un processus séparable vérifiant la condition suivante: il existe une fonction $g(h)$ monotone non décroissante non négative et une fonction $q(c, h)$, $h \geqslant 0$, telles que*

$$\mathsf{P}\{\rho[\xi(t+h), \xi(t)) > Cg(h)\} \leqslant q(C, h) \tag{3}$$

et

$$G = \sum_{n=0}^{\infty} g(2^{-n}T) < \infty, \; Q(C) = \sum_{n=1}^{\infty} 2^n q(C, 2^{-n}T) < \infty. \tag{4}$$

Alors

$$\mathsf{P}\{\sup_{0 \leqslant t' < t'' \leqslant T} \rho(\xi(t'), \xi(t'')) > N\} \leqslant Q\left(\frac{N}{2G}\right) \tag{5}$$

et

$$\mathsf{P}\left\{\sup_{|t'-t''| \leqslant \varepsilon} \rho(\xi(t'), \xi(t'')) > CG\left(\left[\log_2 \frac{T}{2\varepsilon}\right]\right)\right\} \leqslant$$
$$\leqslant Q\left(\left[\log_2 \frac{T}{2\varepsilon}\right], C\right), \tag{6}$$

où

$$G(m) = \sum_{n=m}^{\infty} g(2^{-n}T), \quad Q(mC), = \sum_{n=m}^{\infty} 2^n q(C, 2^{-n}T). \tag{7}$$

Pour démontrer ce lemme il suffit de reprendre, en les simplifiant, les démonstrations des lemmes 3 et 4, § 4. On n'indiquera que les grandes lignes. Soient les événements

$$A_{nk} = \left\{\rho\left(\xi\left(\frac{k+1}{2^n}T\right), \xi\left(\frac{k}{2^n}T\right)\right) \leqslant Cg(2^{-n}T)\right\},$$
$$k = 0, 1, \ldots, 2^n - 1, \; n = 0, 1, 2, \ldots$$

Posons $D_m = \bigcap_{n=m}^{\infty} \bigcap_{k=0}^{2n-1} A_{nk}$. Il vient

$$\mathsf{P}\{\overline{D}_n\} \leqslant Q(n, C).$$

De D_n il suit que $\forall t'$ et t'' de I

$$\rho(\xi(t'), \xi(t'')) \leqslant 2CG;$$

si, outre D_n, on a $0 \leqslant t'' - t' \leqslant 2^{-n}T$, alors $\rho(\xi(t'), \xi(t'')) \leqslant \leqslant 2CG(n)$. En raisonnant comme dans la fin de la démonstration des lemmes du § 4 on obtient ce qu'on avait annoncé. A noter que la continuité stochastique du processus $\xi(t)$ découle des conditions (3) et (4). ∎

THÉORÈME 5. *Si les conditions du lemme* 1 *sont réalisées, le processus* $\xi(t)$ *est continu.*

Comme cas particulier où les conditions (3) et (4) sont réalisées considérons un processus tel que

$$\mathsf{E}\rho^p[\xi(t'), \xi(t'')] \leqslant L\,|t''-t'|^{1+r}, \tag{8}$$

où $p>0$, $r>0$. Posons $g(h)=h^{r'/p}$, où $0<r'<r$. On trouve

$$G\left(\left[\log_2 \frac{T}{2\varepsilon}\right]\right) \leqslant K_1 \varepsilon^{r'/p}$$

et

$$Q\left(\left[\log_2 \frac{T}{2\varepsilon}\right], C\right) \leqslant C^{-p} K_2 \varepsilon^{(r-r')},$$

où K_1 et K_2 sont des constantes. Du théorème 5 il suit le

COROLLAIRE 1. *Si un processus aléatoire séparable* $\xi(t)$ *vérifie la condition* (8), *il est continu.*

Signalons encore une condition plus générale que (8) et telle que (3) et (4) sont réalisées. Soit

$$\mathsf{E}\rho^p[\xi(t), \xi(t+h)] \leqslant \frac{L\,|h|}{|\log_2|h||^{1+r}}, \quad p<r. \tag{9}$$

Si l'on pose $g(h)=|\log_2|h||^{-r'/p}$, où $p<r'<r$, alors

$$G=\sum_{n=0}^{\infty} |\log_2|2^{-n}||^{-r'/p}<\infty, \quad Q(C) \leqslant \frac{LT}{\sum\limits_{n=0}^{\infty} C^p|\log_2|2^{-n}T||^{1+r-r'}}<\infty.$$

COROLLAIRE 2. *Si un processus séparable* $\xi(t)$ *vérifie la relation* (9), *il est continu.*

Processus gaussiens. Appliquons les résultats précédents à un processus gaussien réel séparable unidimensionnel $\xi(t)$, $t \in [0, T]$, de fonction de corrélation $R(t)$ et d'espérance mathématique nulle. La différence $\xi(t+h)-\xi(t)$ a pour variance

$$\sigma^2(t, h) = R(t+h, t+h) - 2R(t, t+h) + R(t, t),$$

donc

$$\mathsf{P}\{|\xi(t+h)-\xi(t)|>Cg(h)\} = \frac{2}{\sqrt{2\pi}} \int_\alpha^\infty e^{-\frac{t^2}{2}}\,dt,$$

où $\alpha = Cg(h)\,\sigma^{-1}(t, h)$. En se servant de l'inégalité

$$\int_\alpha^\infty e^{-\frac{t^2}{2}}\,dt \leqslant \frac{1}{\alpha} e^{-\alpha^2/2}, \tag{10}$$

qui se déduit sans peine par une intégration par parties du premier membre, on trouve

$$\mathsf{P}\{|\xi(t+h)-\xi(t)|>Cg(h)\}\leqslant\frac{2}{\sqrt{2\pi}}\frac{\sigma(t,h)}{Cg(h)}e^{-\frac{C^2g^2(h)}{2\sigma^2(t,h)}}. \quad (11)$$

THÉORÈME 6. *Si un processus gaussien vérifie*

$$\sigma^2(t,h)\leqslant\frac{K}{|\ln|h||^p},\quad p>3, \quad (12)$$

il est continu.

Démonstration. Posons $g(h)=|\ln|h||^{-p'}$, où p' est un nombre tel que $1<p'<\frac{p-1}{2}$. On peut admettre alors que

$$q(C,h)=\frac{K'}{C|\ln|h||^{p/2-p'}}e^{-\frac{C^2}{2K}|\ln|h||^{p-2p'}},$$

et les séries (4) seront convergentes. Ce qui démontre le théorème.

§ 6. Submartingales à argument continu

Les submartingales (supermartingales, martingales) ont été introduites au § 6, chapitre III. Dans le présent paragraphe on verra que les submartingales (et par suite les supermartingales et les martingales) possèdent des modifications continues à droite sous des hypothèses assez larges, et ce sans utiliser les théorèmes d'existence d'une modification séparable.

Signalons tout d'abord que les inégalités établies au § 1, chapitre III, pour les submartingales de suites se transposent sans difficultés aux submartingales $\xi(t)$ définies sur un ensemble dénombrable quelconque $S\subset[0,\infty[$.

Soit $\{\mathfrak{F}_t,\ t\in S\}$ un flot de tribus, $C>0$.

Si $\{\xi(t),\ \mathfrak{F}_t,\ t\in S\}$ est une submartingale, alors

a) $$\mathsf{P}\{\sup_{t\in S}\xi(t)>C\}\leqslant\frac{\sup_{t\in S}\mathsf{E}\xi^+(t)}{C}; \quad (1)$$

b) $$\mathsf{E}\nu([a,b],S)\leqslant\frac{\sup_{t\in S}\mathsf{E}(\xi(t)-b)^+}{C}, \quad (2)$$

où $\nu([a,b],S)$ est le nombre d'intersections de l'intervalle $]a,b]$ de haut en bas par la fonction $f(t)=\xi(t,\omega)$ sur l'ensemble $t\in S$;

c) $$\mathsf{E}\sup_{t\in S}[\xi^+(t)]^p\leqslant q^p\sup_{t\in S}\mathsf{E}[\xi^+(t)]^p \quad (3)$$

si existe $p>1$ tel que $\mathsf{E}[\xi^+(t)]^p<\infty$.

La démonstration est la même que dans le cas où S est la suite $\{1,2,\ldots,n,\ldots\}$.

Soit la submartingale $\{\xi(t), \mathfrak{F}_t, t \in [0, \infty[\}$. Supposons que $\mathfrak{F}_0$ contient tous les ensembles de probabilité nulle. Montrons que sous des hypothèses assez larges il existe une modification du processus $\xi(t)$ dont les réalisations possèdent presque sûrement des limites à gauche et sont continues à droite quel que soit $t > 0$.

Choisissons à cet effet un ensemble S dénombrable partout dense dans $[0, \infty[$ et étudions la restriction de $\xi(t)$ à S. Posons

$$N_{nab} = \{\omega : \nu(]a, b], S \cap [0, n]) = \infty\},$$

$$L'_n = \{\omega : \sup \xi(t) = +\infty, t \in S \cap [0, n]\},$$

$$L''_n = \{\omega : \inf \xi(t) = -\infty, t \in S \cap [0, n]\}.$$

L'inégalité (2) entraîne que $\mathsf{P}(N_{nab}) = 0$ quels que soient n, a, b. Comme

$$\mathsf{P}(L''_n) \leqslant \sum_{m=1}^{\infty} \mathsf{E} \lim_{k\to\infty} \nu(]-k, -m], S \cap [0, n]),$$

il vient en vertu de (2)

$$\mathsf{P}(L''_n) \leqslant \sum_{m=1}^{\infty} \lim_{\substack{k\to\infty \\ k>m}} \frac{\mathsf{E}(\xi(n)+m)^+}{k-m} = 0.$$

Enfin de l'inégalité (1) il suit $\mathsf{P}(L'_n) = 0$. Soit

$$N = \bigcup_{n=1}^{\infty} ((\bigcup_{a,b} N_{nab}) \cup L'_n \cup L''_n),$$

où a et b parcourent l'ensemble des nombres rationnels. Alors $\mathsf{P}(N) = 0$. Il est aisé de voir maintenant que si $\omega \notin N$, en tout point t, $\xi(s)$, $s \in S$, possède la limite $\xi(t-0)$ à gauche et la limite $\xi(t+0)$ à droite. En effet, si en un point t l'une de ces limites n'existait pas, par exemple $\xi(t-0)$, alors $b' = \overline{\lim\limits_{\substack{s\to t \\ s<t}}}\ \xi(s)$ et $a' = \underline{\lim\limits_{\substack{s\to t \\ s<t}}}\ \xi(s)$ seraient finies et inégales entre elles. Or pour tout couple de nombres rationnels (a, b) tels que $a' < a < b < b'$ le nombre d'intersections de l'intervalle $]a, b]$ par le processus $\xi(s)$, $s \in S$, $s < t$, serait infini contrairement à l'hypothèse $\omega \notin N$. Posons pour tout $t \geqslant 0$

$$\eta(t) = \begin{cases} \xi(t+0) & \text{si } \omega \notin N, \\ \xi(t) & \text{si } \omega \in N. \end{cases}$$

Si $\omega \notin N$, la fonction $\eta(t)$ est, pour chaque $t \geqslant 0$, continue à droite et possède une limite à gauche $(\eta(t-0) = \xi(t-0))$.

Soient les tribus $\mathfrak{F}_{t+}$, $t \geqslant 0$, où $\mathfrak{F}_{t+} = \bigcap_{s>t} \mathfrak{F}_s$. De toute évidence, $\{\mathfrak{F}_{t+}, t \geqslant 0\}$ forment un flot de tribus, $\mathfrak{F}_{t+}$ contiennent tous les sous-ensembles de probabilité 0 et $\eta(t)$ est $\mathfrak{F}_{t+}$-mesurable. Prouvons que $\{\eta(t), \mathfrak{F}_{t+}, t \geqslant 0\}$ est une submartingale.

La suite $\xi(s_n)$, $s_1 \geqslant s_2 \geqslant \ldots \geqslant s_n \geqslant \ldots$, $s_n \downarrow s$, étant équi-intégrable et convergente vers $\eta(s)$ dans $\mathscr{L}_1$, il suit que dans l'inégalité

$$\int_B \xi(s_n)\, d\mathsf{P} \leqslant \int_B \xi(t)\, d\mathsf{P},\ s_n < t,\ B \in \mathfrak{F}_{s+},$$

on peut passer à la limite pour $n \to \infty$. On obtient

$$\int_B \eta(s)\, d\mathsf{P} \leqslant \int_B \xi(t)\, d\mathsf{P}.$$

De façon analogue on s'assure que dans le second membre de la dernière inégalité on peut remplacer $\xi(t)$ par $\eta(t)$, c'est-à-dire

$$\int_B \eta(s)\, d\mathsf{P} \leqslant \int_B \eta(t)\, d\mathsf{P}. \tag{4}$$

Ceci montre que $\{\eta(t), \mathfrak{F}_{t+}, t \geqslant 0\}$ est une submartingale.

Théorème 1. *Soient* $\{\xi(t), \mathfrak{F}_{t+}, t > 0\}$ *une submartingale,* $\mathfrak{F}_{t+} = \mathfrak{F}_t$, $a(t) = \mathsf{E}\xi(t)$ *une fonction continue à droite pour tous les* $t \geqslant 0$. *Il existe alors une modification* $\{\eta(t), \mathfrak{F}_t, t \geqslant 0\}$ *du processus* $\xi(t)$ *dont les réalisations sont presque sûrement continues à droite et possèdent des limites à gauche quel que soit* t.

Démonstration. Il suffit de prouver que le processus $\{\eta(t), \mathfrak{F}_t, t \geqslant 0\}$ construit précédemment est une modification du processus donné. Par analogie à (4) on a

$$\int_B \xi(s)\, d\mathsf{P} \leqslant \int_B \eta(t)\, d\mathsf{P} \quad \forall (t, B),\ t \geqslant s,\ B \in \mathfrak{F}_s,$$

ou $\xi(s) \leqslant \mathsf{E}\{\eta(t) \mid \mathfrak{F}_s\}$, $s \leqslant t$. Comme $\eta(t)$ est $\mathfrak{F}_t$-mesurable, on a

$$\xi(t) \leqslant \mathsf{E}\{\eta(t) \mid \mathfrak{F}_t\} = \eta(t) = \xi(t+0) \pmod{\mathsf{P}}.$$

L'équi-intégrabilité de la suite $\xi(t_n)$ pour $t_n \downarrow t$ implique

$$\mathsf{E}(\xi(t+0) - \xi(t)) = \lim_{t_n \downarrow t} \mathsf{E}(\xi(t_n) - \xi(t)) = \lim_{t_n \downarrow t} (a(t_n) - a(t)) = 0,$$

ce qui, compte tenu de l'inégalité précédente, donne

$$\xi(t+0) = \xi(t) \pmod{\mathsf{P}}.$$

Remarque. La condition $\mathfrak{F}_{t+} = \mathfrak{F}_t$ n'a été utilisée que pour affirmer que $\eta(t)$ est une variable aléatoire $\mathfrak{F}_t$-mesurable. Donc le théorème 1 reste en vigueur si l'on remplace cette condition par la

suivante: la variable aléatoire $\xi(t+0) = \lim_{s_n \downarrow t} \xi(s_n)$ est $\mathfrak{F}_t$-mesurable.

COROLLAIRE. *Si* $\xi(t)$, $t \geqslant 0$, *submartingale par rapport à un flot de tribus* $\{\mathfrak{F}_t, t \geqslant 0\}$, *où* $\mathfrak{F}_t$ *est la complétion de la tribu engendrée par les variables aléatoires* $(\xi(s), s \leqslant t)$, *est stochastiquement continue à droite*:

$$\lim_{h \downarrow 0} \mathsf{P}\{|\xi(t)-\xi(t+h)|>\varepsilon\}=0 \ \forall \ (\varepsilon<0, t\geqslant 0),$$

il existe une submartingale $\{\eta(t), \mathfrak{F}_t, t \geqslant 0\}$ *stochastiquement équivalente à un processus* $\xi(t)$ *dont les réalisations sont continues à droite et possèdent des limites à gauche presque sûrement.*

Pour le prouver il faut vérifier que la condition de continuité stochastique entraîne les conditions du théorème 1. Les suites $\xi(t+h)$ convergeant stochastiquement vers $\xi(t)$ pour $h \downarrow 0$, il suit que pour tout t il existe une suite t_n, $t_n \downarrow t$, $t_n \in S$, telle que $\xi(t) = \lim \xi(t_n)$ (mod P). Donc $\xi(t+0)$ est $\mathfrak{F}_t$-mesurable. De l'équi-intégrabilité de la famille $\xi(t_n)$, $n = 1, 2, \ldots$, il suit: $a(t) = \mathsf{E}\xi(t) = \mathsf{E}\lim \xi(t_n) = \lim \mathsf{E}\xi(t_n) = \lim a(t_n)$, c'est-à-dire $a(t+0) = a(t)$. La remarque du théorème 1 entraîne que le processus $\{\eta(t), \mathfrak{F}_t, t > 0\}$ est une submartingale possédant les propriétés requises.

Le théorème 1, § 5, entraîne le

THÉORÈME 2. *Si* $\{\xi(t), t \in [a, b]\}$ *est une submartingale telle que*

$$\lim_{\max_k |t_{n,k+1}-t_{nk}| \to 0} \sum_{k=0}^{n-1} \mathsf{P}\{|\xi(t_{n,k+1})-\xi(t_{nk})|>\varepsilon\}=0,$$

où $a = t_{n0} < t_{n1} < \ldots < t_{nn} = b$, *alors elle possède une modification continue.*

CHAPITRE V

TRANSFORMATIONS LINÉAIRES DE PROCESSUS ALÉATOIRES

Dans ce chapitre on se propose d'étudier des opérations linéaires sur les processus aléatoires, telles la dérivation, l'intégration, les transformations par des équations différentielles ou intégrales. La prédiction de la valeur prise par une fonction aléatoire ou encore la sélection d'une composante parmi les valeurs observées d'une fonction aléatoire représentée par la superposition du signal transmis et du « bruit » sont des problèmes qui entrent dans le cadre de la théorie des transformations linéaires des processus aléatoires.

Ces problèmes sont étudiés dans un espace hilbertien de variables aléatoires, car les solutions qui y sont obtenues sont commodes aux applications. On suppose au lecteur la connaissance des notions élémentaires de la théorie des espaces hilbertiens.

§ 1. Fonctions aléatoires hilbertiennes

Soit $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ un espace probabilisé.

DÉFINITION 1. *On appelle espace hilbertien* $\mathcal{L}_2 = \mathcal{L}_2(\Omega, \mathfrak{S}, \mathsf{P})$ *de variables aléatoires de l'espace probabilisé* $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ *un ensemble de variables aléatoires à valeurs complexes* $\zeta = f(\omega)$, $\omega \in \Omega$, *telles que* $\mathsf{E}\,|\zeta|^2 < \infty$.

Le produit scalaire sur $\mathcal{L}_2$ est défini par

$$(\zeta, \eta) = \mathsf{E}\,\zeta\bar{\eta}, \quad \zeta, \eta \in \mathcal{L}_2.$$

D'après cette définition la norme $\|\zeta\|$ de la variable aléatoire ζ vaut

$$\|\zeta\| = \{\mathsf{E}\,|\zeta|^2\}^{1/2}.$$

Deux variables aléatoires ζ et η sont orthogonales si

$$(\zeta, \eta) = \mathsf{E}\zeta\bar{\eta} = 0.$$

Le carré de la norme $\|\zeta\|^2$ de la variable aléatoire réelle ζ est confondu avec le moment d'ordre deux, $\|\zeta\|^2 = \mathsf{E}\,\zeta^2$, et avec la variance si $\mathsf{E}\,\zeta = 0$. Si ζ et η sont réels et $\mathsf{E}\,\zeta = \mathsf{E}\eta = 0$, leur orthogonalité équivaut à leur non-corrélation.

16—0398

DÉFINITION. *Une fonction aléatoire à valeurs complexes* $\zeta(\theta)$ $(\theta \in \Theta)$ *est hilbertienne si*

$$\mathsf{E} \mid \zeta(\theta) \mid^2 < \infty, \ \theta = \Theta.$$

Une fonction aléatoire hilbertienne peut être traitée comme une fonction définie sur Θ à valeurs dans un espace hilbertien $\mathscr{L}_2$:

$$\theta \to \zeta(\theta) = f(\theta, \omega) \in \mathscr{L}_2.$$

En particulier, si Θ est un intervalle de nombres réels $]a, b[$, toute fonction aléatoire hilbertienne doit être traitée comme une ligne de $\mathscr{L}_2$ d'équation paramétrique $\zeta = \zeta(\theta)$, $\theta \in]a, b[$. Dans ce chapitre on n'envisagera que des fonctions aléatoires hilbertiennes, aussi omettra-t-on souvent de le préciser.

Soit donnée sur Θ une fonction non négative $\psi(\theta)$ prenant des valeurs positives aussi petites que l'on veut.

DÉFINITION. *Une variable aléatoire* $\eta \in \mathscr{L}_2$ *est limite en moyenne quadratique (limite en m.q.) d'une fonction aléatoire hilbertienne* $\zeta(\theta)$ *pour* $\psi(\theta) \to 0$ *si* $\zeta(\theta) \to \eta$ *pour* $\psi(\theta) \to 0$ *au sens de la convergence dans l'espace hilbertien* $\mathscr{L}_2$, *c'est-à-dire si*

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \textit{tel que} \ \mathsf{E} \mid \eta - \zeta(\theta) \mid^2 < \varepsilon^2$$

pour tous les θ *tels que* $0 < \psi(\theta) < \delta$.

En particulier, si Θ est un espace métrique muni de la métrique $r(\theta_1, \theta_2)$, la fonction $\zeta(\theta)$ est par définition *continue en moyenne quadratique* (*continue en m.q.*) *au point* $\theta_0 \in \Theta$ *si*

$$\mathsf{E} \mid \zeta(\theta_1) - \zeta(\theta_0) \mid^2 \to 0 \ \textit{pour} \ r(\theta_1, \theta_0) \to 0. \tag{1}$$

DÉFINITION. *On appelle covariance* $B(\theta_1, \theta_2)$, $(\theta_1, \theta_2) \in \Theta^2$, *d'une fonction aléatoire hilbertienne* $\zeta(\theta)$ *le nombre*

$$B(\theta_1, \theta_2) = \mathsf{E}\zeta(\theta_1)\overline{\zeta(\theta_2)} = (\zeta(\theta_1), \zeta(\theta_2)). \tag{2}$$

LEMME 1. *Pour qu'une fonction aléatoire* $\zeta(\theta)$ *possède une limite en m.q. pour* $\psi(\theta) \to 0$ *il est nécessaire et suffisant qu'existe* $\lim B(\theta_1, \theta_2)$ *pour* $\psi(\theta_1) + \psi(\theta_2) \to 0$. *Si cette condition est réalisée et* $\eta = \underset{\psi(\theta)\to 0}{\text{l. i. m.}} \zeta(\theta)$, *alors*

$$\mathsf{E} \mid \eta \mid^2 = \lim_{\psi(\theta_1)+\psi(\theta_2)\to 0} B(\theta_1, \theta_2). \tag{3}$$

On omettra la démonstration en raison de sa facilité.

COROLLAIRE. *Pour qu'une fonction* $\zeta(\theta)$ *soit continue en m.q. en un point* θ_0 *il est nécessaire et suffisant que soit continue la covariance* $B(\theta_1, \theta_2)$ *au point* (θ_0, θ_0).

La condition nécessaire est une conséquence du lemme 1, la condition suffisante s'obtient par un calcul élémentaire.

REMARQUE. La continuité en m.q. de $\zeta(\theta)$ au point θ_0 entraîne la continuité stochastique de $\zeta(\theta)$ en ce point. En effet, l'inégalité de Tchébychev nous dit que

$$\mathsf{P}\{|\zeta(\theta)-\zeta(\theta_0)|>\varepsilon\}\leqslant\frac{\mathsf{E}\,|\zeta(\theta)-\zeta(\theta_0)|^2}{\varepsilon^2}.$$

La continuité en m.q. de $\zeta(\theta)$ sur Θ n'implique pas la continuité presque sûre des réalisations du processus sur Θ. En effet, pour un processus de Poisson on a

$$\mathsf{E}\,|\,\zeta(t+h)-\zeta(t)\,|^2=\lambda h+(\lambda h)^2,$$

mais les réalisations de $\zeta(t)$ sont discontinues avec une probabilité positive.

L'étude générique des fonctions aléatoires hilbertiennes se ramène à celle de fonctions ordinaires à valeurs dans un espace hilbertien. La covariance, les divers types de convergence, les notions probabilistes spécifiques confèrent des traits particuliers à l'analyse des fonctions aléatoires.

Intégration. Soient $\{\Theta, \mathfrak{A}, m\}$ un espace polonais muni d'une mesure complète σ-finie, $\{\zeta(\theta),\ \theta\in\Theta\}$ une fonction aléatoire hilbertienne. Si la covariance $B(\theta_1, \theta_2)$ est continue m-presque pour tous les θ au point (θ, θ), alors en vertu du lemme 1 et du théorème 1, § 3, chapitre IV, il existe une fonction séparable, mesurable, stochastiquement équivalente à $\zeta(\theta)$. Cette remarque montre à quel point est restrictive l'hypothèse retenue. Le théorème 2 (chapitre IV, § 3) entraîne immédiatement le

THÉORÈME 1. *Si*

$$\int_{\Theta} B(\theta, \theta)\, m(d\theta)<\infty, \tag{4}$$

on a presque sûrement

$$\int_{\Theta} |\zeta(\theta)|^2\, m(d\theta)<\infty$$

et

$$\mathsf{E}\int_{\Theta} |\,\zeta(\theta)\,|^2\, m(d\theta)=\int_{\Theta} B(\theta, \theta)\, m(d\theta). \tag{5}$$

COROLLAIRE. *Soient $f_i(\theta)$, $i=1, 2$, des fonctions de $\mathcal{L}_2(\Theta, \mathfrak{A}, m)$ et soit réalisée la condition* (4). *Les intégrales*

$$\eta_i=\int_{\Theta} f_i(\theta)\,\zeta(\theta)\, m(d\theta)$$

existent presque sûrement et en vertu du théorème de Fubini

$$\mathsf{E}\eta_1\bar{\eta}_2 = \mathsf{E}\int_\Theta\int_\Theta f_1(\theta_1)\overline{f_2(\theta_2)}\zeta(\theta_1)\overline{\zeta(\theta_2)}m(d\theta_1)m(d\theta_2) =$$

$$= \int_\Theta\int_\Theta f_1(\theta_1)B(\theta_1, \theta_2)\overline{f_2(\theta_2)}m(d\theta_1)m(d\theta_2).$$

Faisons quelques remarques au sujet de la définition des intégrales de fonctions aléatoires.

REMARQUE 1. Soit réalisée la condition (4) et $m(\Theta) < \infty$. L'intégrale

$$\int_\Theta \zeta(\theta)m(d\theta), \tag{6}$$

$\zeta(\theta)$ étant une fonction aléatoire mesurable, est définie et presque sûrement finie pour toute réalisation de $\zeta(\theta)$. On peut procéder d'une autre façon pour définir l'intégrale (6). En effet, celle-ci peut être donnée comme la limite en m.q. de sommes intégrales de Lebesgue de $\zeta(\theta)$. On s'assure aisément que cette définition coïncide avec la définition ordinaire. Pour le prouver il suffit de se limiter à des variables aléatoires non négatives. Par définition l'intégrale (6) est la limite pour $n\to\infty$ de

$$\int_\Theta \zeta_n(\theta)m(d\theta),$$

où $\zeta_n(\theta)$ est une suite monotone non décroissante de fonctions aléatoires prenant un nombre fini de valeurs et telles que $\lim \zeta_n(\theta) = \zeta(\theta)\ \forall\theta\in\Theta$. On a

$$\mathsf{E}\left|\int_\Theta \zeta(\theta)m(d\theta) - \int_\Theta \zeta_n(d\theta)m(d\theta)\right|^2 \leqslant \mathsf{E}\int_\Theta |\zeta(\theta)-\zeta_n(\theta)|^2 m(d\theta).$$

Comme $0 \leqslant \zeta(\theta) - \zeta_n(\theta) \leqslant \zeta(\theta)$, il suit du théorème de Lebesgue

$$\int_\Theta \zeta(\theta)m(d\theta) = \text{l.i.m.}\int_\Theta \zeta_n(\theta)m(d\theta).$$

REMARQUE 2. Soit le processus aléatoire $\{\zeta(t),\ t\in[a, b]\}$. L'intégrale

$$\int_a^b \zeta(t)\,dt$$

est souvent définie comme la limite en m.q. des sommes intégrales

$$\sum_{k=1}^n \zeta(t_{nk})\Delta t_{nk},$$

$$\Delta t_{nk} = t_{nk} - t_{n,k-1},\quad a = t_{n0} < t_{n1} < \ldots < t_{nn} = b.$$

Une condition nécessaire et suffisante d'existence de la limite en m. q. de ces sommes est, en vertu du lemme 1, l'existence de la limite de

$$\mathsf{E}\sum_{k=1}^{n}\zeta(t_{nk})\,\Delta t_{nk}\sum_{k=1}^{m}\overline{\zeta(t_{mk})}\,\Delta t_{mk}=\sum_{k=1}^{n}\sum_{r=1}^{m}B(t_{nk},\,t_{mr})\,\Delta t_{nk}\Delta t_{mr}$$

pour $n,\ m\to\infty$, c'est-à-dire l'intégrabilité au sens de Riemann de la fonction $B(t,\ s)$ $(a\leqslant t,\ s\leqslant b)$. Donc la définition donnée est plus restrictive que la définition initiale mais par contre elle n'est pas liée à la mesurabilité du processus. On s'assure immédiatement que l'usage de la dernière définition de l'intégrale conduit (mod P) au même résultat que la définition initiale.

En effet,

$$\mathsf{E}\left|\int_a^b\zeta(t)\,dt-\sum_{k=1}^{n}\zeta(t_{nk})\,\Delta t_{nk}\right|^2=$$

$$=\sum_{k=1}^{n}\sum_{r=1}^{n}\int_{t_{n,\,k-1}}^{t_{nk}}\int_{t_{n,\,r-1}}^{t_{nr}}[B(t,\,s)-B(t,\,t_{nr})-B(t_{nk},\,s)+$$

$$+B(t_{nk},\,t_{nr})]\,dt\,ds\leqslant 2\sum_{k=1}^{n}\sum_{r=1}^{n}\Omega_{nkr}\to 0,$$

où Ω_{nkr} est l'oscillation de la fonction $B(t,\ s)$ dans le rectangle $t_{n,k-1}\leqslant t\leqslant t_{nk},\ t_{n,r-1}\leqslant s\leqslant t_{nr}$.

Remarque 3. Par intégrale impropre en m.q.

$$\int_{-\infty}^{\infty}\zeta(t)\,dt\ \left(\text{ou}\int_a^{\infty}\zeta(t)\,dt\right) \tag{7}$$

on conviendra d'entendre

$$\underset{N\to\infty}{\text{l.i.m.}}\int_{-N}^{N}\zeta(t)\,dt\ \left(\text{l.i.m.}\int_a^{N}\zeta(t)\,dt\right).$$

Pour que ces intégrales existent, il est nécessaire et suffisant, en vertu du lemme 1, qu'existent les limites

$$\lim_{N,\,N'\to\infty}\int_{-N}^{N}\int_{-N'}^{N'}B(t,\,s)\,dt\,ds\ \left(\lim_{N,\,N'\to\infty}\int_a^{N}\int_a^{N'}B(t,\,s)\,dt\,ds\right).$$

Cette définition des intégrales impropres est dans certains cas plus large que la définition des intégrales (7) comme intégrales au sens de Lebesgue de fonctions $\zeta(t)$ à ω fixe.

Loi des grands nombres. Soit $\{\zeta(t),\ t \geqslant 0\}$ un processus hilbertien mesurable à covariance intégrable sur tout intervalle fini. On dira que $\zeta(t)$ vérifie la loi des grands nombres si dans un certain sens

$$\frac{1}{T}\int_0^T \zeta(t)\,dt \to c \text{ pour } T \to \infty.$$

Le lemme 1 nous apprend qu'une condition nécessaire et suffisante de l'existence de la moyenne

$$\underset{T\to\infty}{\text{l.i.m.}}\frac{1}{T}\int_0^T \zeta(t)\,dt$$

est l'existence de la limite

$$\lim_{T,\,T'\to\infty} \mathsf{E}\frac{1}{T}\int_0^T \zeta(t)\,dt\,\frac{1}{T'}\int_0^{T'} \overline{\zeta(t)}\,dt = \lim_{T,\,T'\to\infty} \frac{1}{TT'}\int_0^T\int_0^{T'} B(t,\,s)\,dt\,ds.$$

D'autre part, pour que

$$\underset{T\to\infty}{\text{l.i.m.}}\left\{\frac{1}{T}\int_0^T \zeta(t)\,dt - \frac{1}{T}\int_0^T \mathsf{E}\zeta(t)\,dt\right\} = 0$$

ait lieu, il est nécessaire et suffisant que

$$\lim_{T,\,T'\to\infty} \frac{1}{TT'}\int_0^T\int_0^{T'} R(t,\,s)\,dt\,ds = 0, \tag{8}$$

où $R(t,\,s)$ est la fonction de corrélation du processus.

Il est aisé de remarquer que

$$\left|\int_0^T\int_0^T R(t,\,s)\,dt\,ds\right|^2 \leqslant \int_0^T\int_0^T R(t,\,s)\,dt\,ds \int_0^{T'}\int_0^{T'} R(t,\,s)\,dt\,ds.$$

Donc (8) a lieu si et seulement si

$$\lim_{T\to 0}\frac{1}{T^2}\int_0^T\int_0^T R(t,\,s)\,dt\,ds = 0. \tag{9}$$

Dans le cas d'un processus stationnaire au sens large, $R(t,\,s) = R(t-s)$. Comme

$$\frac{1}{T^2}\int_0^T\int_0^T R(t-s)\,dt\,ds = \frac{1}{T}\int_{-T}^T R(t)\left(1-\frac{|t|}{T}\right)dt,$$

il suit le

THÉORÈME 2. *Si* $\zeta(t)$ *est un processus stationnaire au sens large, pour que*

$$\underset{T\to\infty}{\text{l.i.m.}}\frac{1}{T}\int_0^T \zeta(t)\,dt = \mathsf{E}\zeta(t) \tag{10}$$

ait lieu il est nécessaire et suffisant que

$$\lim_{T\to\infty}\frac{1}{T}\int_{-T}^{T} R(t)\left(1-\frac{|t|}{T}\right)dt = 0. \tag{11}$$

La condition (11) est réalisée en particulier si la valeur moyenne de la fonction de corrélation est nulle:

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} R(s)\,ds = 0.$$

Exprimons la condition (11) au moyen de la fonction spectrale du processus. On a

$$\frac{1}{T}\int_{-T}^{T} R(t)\left(1-\frac{|t|}{T}\right)dt = \int_{-\infty}^{\infty} F(du)\frac{1}{T}\int_{-T}^{T} e^{itu}\left(1-\frac{|t|}{T}\right)dt,$$

d'où

$$\frac{1}{T}\int_{-T}^{T} R(t)\left(1-\frac{|t|}{T}\right)dt = \int_{-\infty}^{\infty}\frac{2(1-\cos Tu)}{T^2u^2}F(du) =$$

$$= F\{0\} + \int_{-\infty}^{\infty}\frac{2(1-\cos Tu)}{T^2u^2}\widetilde{F}(du),$$

où $\widetilde{F}(A) = F(A \setminus \{0\})$, $\{0\}$ étant l'ensemble constitué du seul point $u = 0$. On voit sans peine que la dernière intégrale tend vers 0 pour $T \to \infty$. Donc

$$\lim_{T\to\infty}\frac{1}{T}\int_{-T}^{T} R(t)\left(1-\frac{|t|}{T}\right)dt = F\{0\}. \tag{12}$$

D'où il suit le

THÉORÈME 3. *Une condition nécessaire et suffisante pour que l'égalité* (10) *ait lieu pour un processus stationnaire au sens large est que sa fonction spectrale soit continue au point* $u = 0$.

Dérivation. Soit $\{\zeta(t),\ t \in\,]a,\ b[\}$, $-\infty \leqslant a < b \leqslant \infty$, un processus aléatoire hilbertien.

DÉFINITION. *On dit que le processus aléatoire* $\{\zeta(t),\ t \in]a,\ b[\}$ *est dérivable en m.q. en* t_0 *si existe*

$$\zeta'(t_0) = \underset{n\to 0}{\mathrm{l.i.m}} \frac{\zeta(t_0+h)-\zeta(t_0)}{h},\quad t_0,\ t_0+h \in]a,\ b[.$$

Le nombre aléatoire $\zeta'(t_0)$ *est appelé dérivée en m.q. du processus aléatoire au point* t_0.

On établit sans peine les conditions nécessaires et suffisantes de dérivabilité en m.q. d'un processus aléatoire. Comme

$$\mathsf{E}\,\frac{\zeta(t_0+h)-\zeta(t_0)}{h}\cdot\overline{\frac{\zeta(t_0+h_1)-\zeta(t_0)}{h_1}} =$$

$$= \frac{1}{h_1 h}\{B(t_0+h,\ t_0+h_1) - B(t_0,\ t_0+h_1) - B(t_0+h,\ t_0) + B(t_0,\ t_0)\},$$

il suit du lemme 1 qu'une condition nécessaire et suffisante de dérivabilité en m.q. du processus $\zeta(t)$ au point t_0 est l'existence de la dérivée mixte généralisée:

$$\left.\frac{\partial^2 B(t,\ t')}{\partial t\,\partial t'}\right|_{t=t'=t_0} =$$

$$= \lim_{h,\ h_1\to 0} \frac{B(t_0+h,\ t_0+h_1) - B(t_0,\ t_0+h_1) - B(t_0+h,\ t_0) + B(t_0,\ t_0)}{hh_1}.$$

La dérivabilité en m.q. au point t et l'inégalité

$$\left|\mathsf{E}\left(\zeta'(t) - \frac{\zeta(t+h)-\zeta(t)}{h}\right)\right| \leqslant \left\{\mathsf{E}\left|\zeta'(t) - \frac{\zeta(t+h)-\zeta(t)}{h}\right|^2\right\}^{1/2}$$

entraînent

$$\mathsf{E}\zeta'(t) = \frac{d}{dt}\,\mathsf{E}\zeta(t) \tag{13}$$

et de plus la dérivée à droite existe.

Si le processus $\zeta(t)$ est dérivable en m. q. en tout point de l'intervalle $]a,\ b[$, la dérivée $\zeta'(t)$ constitue un processus aléatoire hilbertien sur $]a,\ b[$.

THÉORÈME 4. *Soit* $\{\zeta(t),\ t \in]a,\ b[\}$ *un processus aléatoire hilbertien. Si la dérivée généralisée*

$$\left.\frac{\partial^2 B(t,\ t')}{\partial t\,\partial t'}\right|_{t=t'}$$

existe pour toute valeur de $t \in]a,\ b[$, *le processus* $\zeta(t)$ *est dérivable en m.q. sur* $]a,\ b[$ *et*

$$B_{\zeta'\zeta'}(t,\ t') = \frac{\partial^2 B(t,\ t')}{\partial t\,\partial t'}, \tag{14}$$

$$B_{\zeta'\zeta}(t,\ t') = \frac{\partial B(t,\ t')}{\partial t}, \tag{15}$$

où $B_{\zeta'\zeta'}(t, t') = \mathsf{E}\zeta'(t)\overline{\zeta'(t')}$ *est la covariance du processus* $\zeta'(t)$ *et* $B_{\zeta'\zeta}(t, t') = \mathsf{E}\zeta'(t)\,\zeta(t')$ *la covariance des processus* $\zeta'(t)$ *et* $\zeta(t)$.

Seules les formules (14) et (15) sont à démontrer. On a

$$B_{\zeta'\zeta}(t, t') = \mathsf{E}\overline{\zeta(t')}\,\zeta'(t) =$$

$$= \lim_{h\to 0} \mathsf{E}\overline{\zeta(t')}\left(\frac{\zeta(t+h)-\zeta(t)}{h}\right) = \lim_{h\to 0}\frac{B(t+h, t')-B(t, t')}{h}.$$

Donc la dérivée $\dfrac{\partial B(t, t')}{\partial t}$ existe et la covariance des proces us $\zeta'(t)$ et $\zeta(t)$ est définie par la formule (15). D'autre part,

$$B_{\zeta'\zeta'}(t, t') = \lim_{h,\,h'\to 0} \mathsf{E}\frac{(\zeta(t'+h')-\zeta(t'))\,(\zeta(t+h)-\zeta(t))}{h'h} =$$

$$= \lim_{h,\,h'\to 0}\frac{B(t+h, t'+h')-B(t, t'+h')-B(t+h, t')+B(t, t')}{hh'}.$$

D'où suit l'existence de la dérivée seconde généralisée

$$\frac{\partial^2 B(t, t')}{\partial t\,\partial t'}$$

(dans l'énoncé on n'a admis que l'existence de la dérivée seconde généralisée pour $t = t'$) et la formule (14). ■

Si le processus $\zeta(t)$ est stationnaire au sens large, alors $B(t, t') = R(t - t')$ et le théorème 4 entraîne le

COROLLAIRE 1. *Pour qu'un processus stationnaire au sens large* $\zeta(t)$ $(t \in T)$ *soit dérivable en m.q. il est nécessaire et suffisant qu'existe la dérivée seconde généralisée de la fonction de corrélation* $R(t)$ *pour* $t = 0$. *Si cette condition est réalisée, il existe alors la dérivée généralisée* $\dfrac{d^2R(t)}{dt^2}$ *et*

$$R_{\zeta'\zeta'}(t_0, t_0+t) = -\frac{d^2R(t)}{dt^2},$$

$$R_{\zeta'\zeta}(t_0+t, t_0) = R_{\zeta'\zeta}(t) = \frac{dR(t)}{dt}.$$

Des résultats analogues ont lieu pour les dérivées en m. q. d'ordre supérieur.

COROLLAIRE 2. *Si* $\zeta(t)$, $t \in\,]-\infty, \infty[$, *est un processus stationnaire au sens large et*

$$\int_{-\infty}^{\infty} u^2 F(du) < \infty,$$

où F est la mesure spectrale du processus, alors le processus $\zeta(t)$ est dérivable en m.q., le processus $(\zeta'(t), \zeta(t))$ est stationnaire au sens large et sa fonction matricielle de corrélation $R(t)$ est de la forme

$$\left\| \begin{matrix} \int_{-\infty}^{\infty} e^{itu} u^2 F(du) & \int_{-\infty}^{\infty} e^{itu} iu F(du) \\ -\int_{-\infty}^{\infty} e^{-itu} iu F(du) & \int_{-\infty}^{\infty} e^{itu} F(du) \end{matrix} \right\|.$$

Développement d'un processus aléatoire en séries orthogonales. Soit $\{\zeta(t),\ t \in [a, b]\}$ un processus hilbertien mesurable continu en m.q. Sa covariance $B(t_1, t_2)$ est un noyau continu défini non négatif sur le carré $[a, b] \times [a, b]$. La théorie des équations intégrales nous dit que le noyau $B(t_1, t_2)$ est développable en une série uniformément convergente sur ses fonctions propres $\varphi_n(t)$;

$$B(t_1, t_2) = \sum_{n=1}^{\infty} \lambda_n \varphi_n(t_1) \overline{\varphi_n(t_2)},$$

où

$$\lambda_n \varphi_n(t) = \int_a^b B(t, s) \varphi_n(s)\, ds, \qquad \int_a^b \varphi_n(t) \overline{\varphi_m(t)}\, dt = \delta_{nm},$$

les valeurs propres λ_n étant positives.

Posons

$$\xi_n = \int_a^b \zeta(t) \overline{\varphi_n(t)}\, dt.$$

Cette intégrale existe (théorème 1) et du corollaire du théorème 1 il vient

$$\mathsf{E}\xi_n \bar{\xi}_m = \int_a^b \int_a^b B(t, s) \overline{\varphi_n(t)} \varphi_m(s)\, dt\, ds = \lambda_n \delta_{nm},$$

c'est-à-dire la suite de variables aléatoires ξ_n $(n = 1, 2, \ldots)$ est orthogonale. On a ensuite

$$\mathsf{E}\zeta(t) \bar{\xi}_n = \int_a^b B(t, s) \varphi_n(s)\, ds = \lambda_n \varphi_n(t).$$

D'où il suit que

$$\mathsf{E} \left| \zeta(t) - \sum_{k=1}^{n} \xi_k \varphi_k(t) \right|^2 =$$

$$= B(t, t) - 2\sum_{k=1}^{n} \overline{\varphi_k(t)}\, \mathsf{E}\zeta(t)\bar{\xi}_k + \sum_{k=1}^{n} \lambda_k |\varphi_k(t)|^2 =$$

$$= B(t, t) - \sum_{k=1}^{n} \lambda_k |\varphi_k(t)|^2 \to 0$$

pour $n \to \infty$ uniformément en t d'après le théorème de Dini.

THÉORÈME 5. *Tout processus hilbertien mesurable continu en m.q.* $\zeta(t)$, $t \in [a, b]$, *se développe en la série*

$$\zeta(t) = \sum_{k=1}^{\infty} \xi_k \varphi_k(t) \tag{16}$$

convergente dans $\mathscr{L}_2$ *pour tout* $t \in [a, b]$. ξ_k *est une suite orthogonale de variables aléatoires,* $\mathsf{E}\,|\xi_k|^2 = \lambda_k$, λ_k *sont les valeurs propres,* $\varphi_k(t)$ *les fonctions propres de la covariance du processus.*

REMARQUE. Si le processus $\zeta(t)$ est gaussien, sa dérivée en m.q. et les intégrales de la forme $\int_a^b f(t)\,\zeta(t)\,dt$ sont des variables aléatoires gaussiennes. Donc si $\zeta(t)$ est un processus gaussien réel et $\mathsf{E}\,\zeta(t) = 0$, les coefficients ξ_k de la série (16) sont des variables gaussiennes indépendantes et la série (16) est presque sûrement convergente pour tout t.

En effet, les variables ξ_k sont indépendantes car orthogonales et gaussiennes. Pour que la série (16) soit presque sûrement convergente il suffit que la série $\sum_{k=1}^{\infty} \mathsf{E}\,(\xi_k\varphi_k(t))^2 = \sum_{k=1}^{\infty} \lambda_k\,|\varphi_k(t)|^2$ soit convergente. Or on a vu que cette série était convergente (et sa somme égale à $B(t, t)$).

Considérons à titre d'exemple le développement en série orthogonale d'un processus de mouvement brownien sur l'intervalle [0, 1]. On admet que $\zeta(0) = 0$, $\mathsf{E}\zeta(t) = 0$, $\operatorname{Var}\zeta(t) = t$, $B(t, s) = \mathsf{E}\zeta(t)\,\zeta(s) = \min(t, s)$. Le calcul des valeurs et des fonctions propres du noyau $B(t, s)$ est immédiat. De l'égalité

$$\lambda_n\varphi_n(t) = \int_0^1 \min(t, s)\,\varphi_n(s)\,ds = \int_0^1 s\varphi_n(s)\,ds + \int_t^1 t\varphi_n(s)\,ds,$$

il suit que $\varphi_n(0) = 0$. En dérivant par rapport à t on trouve

$$\lambda_n\varphi_n'(t) = \int_t^1 \varphi_n(s)\,ds,$$

d'où $\varphi_n'(1) = 0$. Une nouvelle dérivation nous conduit à l'équation

$$\lambda_n\varphi_n''(t) = -\varphi_n(t).$$

Ses solutions normées vérifiant les conditions aux limites $\varphi_n(0) = 0$, $\varphi'_n(1) = 0$ se notent

$$\varphi_n(t) = \sqrt{2}\sin\left(n+\frac{1}{2}\right)\pi t, \quad \lambda_n^{-1} = \left(n+\frac{1}{2}\right)^2\pi^2, \quad n = 0, 1, \ldots,$$

donc

$$\zeta(t) = \sqrt{2}\sum_{n=0}^{\infty}\xi_n\frac{\sin\left(n+\frac{1}{2}\right)\pi t}{\left(n+\frac{1}{2}\right)\pi}, \tag{17}$$

où ξ_n est une suite de variables aléatoires gaussiennes indépendantes de paramètres (0, 1). Cette série est presque sûrement convergente pour t fixe.

Le processus de mouvement brownien se décompose encore de la façon suivante. Soit $\xi(t) = \zeta(t) - t\zeta(1)$. Alors $\xi(t)$ est un processus gaussien de covariance $B_1(t, s) = \min(t, s) - ts$ et $\mathsf{E}\xi(t) = 0$. On calcule les valeurs et les fonctions propres du noyau $B_1(t, s)$ comme dans le cas précédent. On retrouve de nouveau l'équation

$$\lambda_n\varphi''_n(t) = -\varphi_n(t)$$

avec les conditions aux limites $\varphi_n(0) = \varphi_n(1) = 0$, dont les solutions s'écrivent

$$\varphi_n(t) = \sqrt{2}\sin n\pi t, \quad \lambda_n^{-1} = n^2\pi^2, \quad n = 1, 2, \ldots$$

Donc

$$\xi(t) = \zeta(t) - t\zeta(1) = \sqrt{2}\sum_{n=1}^{\infty}\xi_n\frac{\sin n\pi t}{n\pi},$$

où ξ_n $(n = 1, 2, \ldots)$ est une suite normée de variables aléatoires gaussiennes indépendantes telles que

$$\xi_n = \sqrt{2}\int_0^1 \xi(t)\sin n\pi t\,dt.$$

Comme $\mathsf{E}\zeta(1) = 1$, $\mathsf{E}\zeta^2(1) = 1$,

$$\mathsf{E}\xi_n\zeta(1) = \sqrt{2}\int_0^1 \mathsf{E}(\zeta(t) - t\zeta(1))\,\zeta(1)\sin n\pi t\,dt = 0,$$

en posant $\xi_0 = \zeta(1)$ on obtient

$$\zeta(t) = t\xi_0 + \sqrt{2}\sum_{n=1}^{\infty}\xi_n\frac{\sin n\pi t}{n\pi}, \tag{18}$$

où $\xi_0, \xi_1, \ldots, \xi_n, \ldots$ sont indépendantes et possèdent une répartition normale de paramètres (0, 1). La série (18) a le même caractère de convergence que la série (17).

§ 2. Mesures et intégrales stochastiques

Les intégrales de la forme

$$\int_a^b f(t)\, d\zeta(t), \tag{1}$$

où $f(t)$ est une fonction non aléatoire donnée et $\zeta(t)$ un processus aléatoire, jouent un rôle important dans de nombreux problèmes. Les réalisations du processus $\zeta(t)$ sont généralement des fonctions à variation non bornée et l'intégrale (1) n'est pas une intégrale de Stieltjes ou de Lebesgue-Stieltjes pour presque toutes les réalisations de $\zeta(t)$. Néanmoins on peut définir l'intégrale (1) de telle sorte qu'elle possède les propriétés de l'intégrale ordinaire.

Dans le présent paragraphe on définit et étudie les propriétés de l'intégrale par rapport à une masse aléatoire. De telles intégrales sont dites *stochastiques.*

Soient $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ un espace probabilisé, $\mathcal{L}_2 = \mathcal{L}_2(\Omega, \mathfrak{S}, \mathsf{P})$, E un ensemble et $\mathfrak{M}$ un semi-anneau de parties de E. Supposons qu'à tout $\Delta \in \mathfrak{M}$ est associée une variable aléatoire $\zeta(\Delta)$ à valeurs complexes telle que

1) $\zeta(\Delta) \in \mathcal{L}_2$, $\zeta(\varnothing) = 0$;
2) $\zeta(\Delta_1 \cup \Delta_2) = \zeta(\Delta_1) + \zeta(\Delta_2) \pmod{\mathsf{P}}$ si $\Delta_1 \cap \Delta_2 = \varnothing$;
3) $\mathsf{E}\zeta(\Delta_1)\overline{\zeta(\Delta_2)} = m(\Delta_1 \cap \Delta_2)$,

où $m(\Delta)$ est une fonction d'ensembles sur $\mathfrak{M}$.

DÉFINITION 1. *Toute famille de variables aléatoires* $\{\zeta(\Delta), \Delta \in \mathfrak{M}\}$ *vérifiant les conditions* 1) *à* 3) *est dite mesure stochastique orthogonale élémentaire et* $m(\Delta)$, *sa fonction structurelle.*

L'orthogonalité de la mesure stochastique est exprimée par la condition 3): si $\Delta_1 \cap \Delta_2 = \varnothing$, les variables $\zeta(\Delta_1)$ et $\zeta(\Delta_2)$ sont orthogonales.

Par définition $m(\Delta)$ n'est pas négative:

$$m(\Delta) = \mathsf{E}\,|\zeta(\Delta)|^2 \geqslant 0, \quad m(\varnothing) = 0,$$

et est additive; si $\Delta_1 \cap \Delta_2 = \varnothing$, alors

$$m(\Delta_1 \cup \Delta_2) = \mathsf{E}\,|\zeta(\Delta_1) + \zeta(\Delta_2)|^2 =$$
$$= m(\Delta_1) + m(\Delta_2) + 2m(\Delta_1 \cap \Delta_2) = m(\Delta_1) + m(\Delta_2).$$

Donc $m(\Delta)$ est une prémesure (chapitre II, § 2) sur $\mathfrak{M}$.

Appelons $\mathcal{L}_0\{\mathfrak{M}\}$ la classe des fonctions simples $f(x)$:

$$f(x) = \sum_{k=1}^{n} c_k \chi_{\Delta_k}(x), \quad \Delta_k \in \mathfrak{M}, \quad k = 1, 2, \ldots, n, \tag{2}$$

où n est un nombre quelconque et $\chi_A(x)$ l'indicateur de l'ensemble A.

Définissons l'intégrale stochastique d'une fonction $f(x) \in \mathcal{L}_0\{\mathfrak{M}\}$ par rapport à la mesure élémentaire stochastique $\zeta(\Delta)$ par

$$\eta = \int f(x)\,\zeta(dx) = \sum_{k=1}^{n} c_k \zeta(\Delta_k). \tag{3}$$

Comme $\mathfrak{M}$ est un semi-anneau, tout couple de fonctions de $\mathcal{L}_0\{\mathfrak{M}\}$ est représentable par des combinaisons linéaires d'indicateurs de mêmes ensembles de $\mathfrak{M}$. Donc si $f, g \in \mathcal{L}_0\{\mathfrak{M}\}$, on admettra que $f(x)$ est définie par la formule (2) et $g(x) = \sum_{k=1}^{n} d_k \chi_{\Delta_k}(x)$, et de plus $\Delta_k \cap \Delta_r = \varnothing$ pour $k \neq r$.

L'orthogonalité de $\zeta(\Delta)$ donne

$$\mathsf{E}\left(\int f(x)\,\zeta(dx) \int \overline{g(x)\,\zeta(dx)}\right) = \sum_{k=1}^{n} c_k \bar{d}_k. \tag{4}$$

Supposons que la prémesure m vérifie la condition de semi-additivité et par suite est prolongeable à une mesure complète $\{\mathsf{E}, \mathfrak{B}, m\}$. Alors $\mathcal{L}_0\{\mathfrak{M}\}$ est un sous-ensemble vectoriel de l'espace hilbertien $\mathcal{L}_2\{m\} = \mathcal{L}_2\{\mathsf{E}, \mathfrak{B}, m\}$. Appelons $\mathcal{L}_2\{\mathfrak{M}\}$ la fermeture de $\mathcal{L}_0\{\mathfrak{M}\}$ dans $\mathcal{L}_2\{m\}$.

L'égalité (4) se transcrit

$$\mathsf{E}\int f(x)\,\zeta(dx) \int \overline{g(x)\,\zeta(dx)} = \int f(x)\,\overline{g(x)}\,m(dx) \tag{5}$$

pour tout couple de fonctions $f(x)$, $g(x) \in \mathcal{L}_0\{m\}$.

Soit maintenant l'enveloppe linéaire $\mathcal{L}_0\{\zeta\}$ de la famille de variables aléatoires $\{\zeta(\Delta),\ \Delta \in \mathfrak{M}\}$, c'est-à-dire l'ensemble de toutes les variables aléatoires représentables par (3); soit d'autre part $\mathcal{L}_2\{\zeta\}$ la fermeture de $\mathcal{L}_0\{\zeta\}$ dans l'espace hilbertien des variables aléatoires $\mathcal{L}_2\{\Omega, \mathfrak{S}, \mathsf{P}\}$. On remarquera que la relation (3) établit une isométrie $\eta = \psi(f)$ entre $\mathcal{L}_0\{\mathfrak{M}\}$ et $\mathcal{L}_0\{\zeta\}$. Cette isométrie est prolongeable à une isométrie ψ entre $\mathcal{L}_2\{\mathfrak{M}\}$ et $\mathcal{L}_2\{\zeta\}$. Si $\eta = \psi(f)$, $f \in \mathcal{L}_2\{\mathfrak{M}\}$, on pose par définition

$$\eta = \psi(f) = \int f(x)\,\zeta(dx) \tag{6}$$

et on appelle la variable aléatoire η *intégrale stochastique* de la fonction $f(x)$ par rapport à ζ. D'où suit le

THÉORÈME 1.

a) *L'intégrale stochastique de toute fonction simple* (2) *est définie par la formule* (3);

b) *l'égalité* (5) *est vraie quelles que soient* $f(x), g(x) \in \mathcal{L}_2\{\mathsf{E}, \mathfrak{B}, m\}$;

c) $$\int [\alpha f(x) + \beta g(x)]\,\zeta(dx) = \alpha \int f(x)\,\zeta(dx) + \beta \int g(x)\,\zeta(dx): \tag{7}$$

d) *pour toute suite de fonctions* $f^{(n)}(x) \in \mathcal{L}_2\{\mathsf{E}, \mathfrak{B}, m\}$ *telle que*

$$\int |f(x) - f^{(n)}(x)|^2 m(dx) \to 0, \quad n \to \infty, \tag{8}$$

on a

$$\int f(x)\zeta(dx) = \text{l.i.m.} \int f^{(n)}(x)\zeta(dx).$$

Remarque. En particulier, si $f^{(n)}(x)$ sont des fonctions simples

$$f^{(n)}(x) = \sum_{k=1}^{m_n} c_k^{(n)}\chi_{\Delta_k^{(n)}}(x), \quad \Delta_k^{(n)} \in \mathfrak{M}, \quad n = 1, 2, \ldots,$$

et (8) est réalisé, alors

$$\int f(x)\zeta(dx) = \text{l.i.m.} \sum_{k=1}^{m_n} c_k^{(n)}\zeta(\Delta_k^{(n)}).$$

L'existence d'une suite de fonctions simples approximant une fonction quelconque $f(x) \in \mathcal{L}_2\{\mathsf{E}, \mathfrak{B}, m\}$ est la conséquence de théorèmes généraux de la théorie de la mesure. Donc l'intégrale stochastique peut être traitée comme la limite en m.q. de sommes intégrales appropriées.

Appelons $\mathfrak{B}_0$ la classe des ensembles $A \in \mathfrak{B}$ pour lesquels $m(\Delta) < \infty$. Soit $\tilde{\zeta}(A)$ fonction aléatoire d'ensembles définie comme suit:

$$\tilde{\zeta}(A) = \int \chi_A(x)\zeta(dx) = \int_A \zeta(dx). \tag{9}$$

Cette fonction possède les propriétés suivantes:

a) $\tilde{\zeta}(A)$ est définie sur la classe d'ensembles $\mathfrak{B}_0$;

b) si $A_n \in \mathfrak{B}_0$, $n = 0, 1, 2, \ldots$, $A_0 = \bigcup_{n=1}^{\infty} A_n$, $A_k \cap A_r = \varnothing$ pour $k \neq r$, $k > 0$, $r > 0$, alors $\tilde{\zeta}(A_0) = \sum_{n=1}^{\infty} \tilde{\zeta}(A_n)$ au sens de la convergence en m.q.;

c) $\mathsf{E}\tilde{\zeta}(A)\overline{\tilde{\zeta}(B)} = m(A \cap B)$, $A, B \in \mathfrak{B}_0$;

d) $\tilde{\zeta}(\Delta) = \zeta(\Delta)$ pour $\Delta \in \mathfrak{M}$.

Définition 2. *On appelle mesure orthogonale stochastique toute fonction aléatoire d'ensembles* $\tilde{\zeta}$ *vérifiant les conditions* a), b), c).

La propriété d) exprime le fait que $\tilde{\zeta}(\Delta)$ est le prolongement de la mesure stochastique élémentaire $\zeta(\Delta)$. On a donc le

Théorème 2. *Toute mesure stochastique élémentaire* $\zeta(\Delta)$ *dont la fonction structurelle est semi-additive est prolongeable à une mesure stochastique* $\tilde{\zeta}(\Delta)$.

REMARQUE. Comme $\mathscr{L}_2\{\zeta\} = \mathscr{L}_2\{\tilde{\zeta}\}$, on a

$$\int f(x)\,\zeta(dx) = \int f(x)\,\tilde{\zeta}(dx).$$

En vertu de cette égalité on conviendra dans la suite d'identifier les intégrales par rapport à la mesure orthogonale élémentaire $\zeta(\Delta)$ dont les fonctions structurelles sont semi-additives aux intégrales stochastiques par rapport à la mesure stochastique $\tilde{\zeta}$ définie par (9).

Faisons quelques remarques sur la définition de l'intégrale stochastique sur un segment de droite. Soit $\xi(t)$ $(a \leqslant t < b)$ un processus à accroissements orthogonaux, c'est-à-dire

$$\mathsf{E}\,(\xi(t_2) - \xi(t_1))\,\overline{(\xi(t_4) - \xi(t_3))} = 0$$

quels que soient $t_i \in [a, b[$, $t_1 < t_2 < t_3 < t_4$, continu à gauche en m.q. :

$$\mathsf{E}\,|\,\xi(t) - \xi(s)\,|^2 \to 0 \text{ pour } s \uparrow t.$$

Posons

$$F(t) = \mathsf{E}\,|\,\xi(t) - \xi(a)\,|^2.$$

De l'orthogonalité des accroissements du processus $\xi(t)$ il suit pour $t_2 > t_1$

$$F(t_2) = \mathsf{E}\,|\,\xi(t_2) - \xi(t_1) + \xi(t_1) - \xi(a)\,|^2 = F(t_1) + \mathsf{E}\,|\,\xi(t_2) - \xi(t_1)\,|^2,$$

d'où

$$F(t_2) \geqslant F(t_1) \quad \text{et} \quad F(t) = \lim_{s \uparrow t} F(s).$$

Par suite $F(t)$ est une fonction monotone non décroissante continue à gauche. Soit $\mathfrak{M}$ la classe des intervalles ouverts

$$\Delta = [t_1, t_2[,\ a \leqslant t_1 < t_2 \leqslant b, \quad \zeta([t_1, t_2[) = \xi(t_2) - \xi(t_1),$$

$m([t_1, t_2[) = F(t_2) - F(t_1)$. Alors $\mathfrak{M}$ est un semi-anneau d'ensembles

$$\mathsf{E}\,\zeta(\Delta_1)\,\overline{\zeta(\Delta_2)} = m(\Delta_1 \cap \Delta_2),$$

$\zeta(\Delta)$ est une mesure stochastique orthogonale élémentaire dont la fonction structurelle est prolongeable à une mesure. On peut donc définir l'intégrale stochastique de Stieltjes en posant

$$\int_a^b f(t)\,d\xi(t) = \int_a^b f(t)\,\zeta(dt),$$

où $\xi(t)$ est un processus à accroissements orthogonaux. Cette intégrale existe pour toute fonction borélienne $f(t)$, $t \in [a, b[$ telle que

$$\int_a^b |\,f(t)\,|^2\,F(dt) < \infty,$$

où $F(A)$ est une mesure associée à la fonction monotone $F(t)$. On

définit de façon analogue l'intégrale stochastique sur la droite $]-\infty, \infty[$ tout entière.

Démontrons quelques propositions sur les intégrales stochastiques.

Soient $\zeta(\cdot)$ une mesure stochastique orthogonale dont la fonction structurelle m est une mesure complète sur $\{E, \mathfrak{B}\}$, et $g(x) \in \mathcal{L}_2(m)$. Posons

$$\lambda(A) = \int \chi_A(x)\, g(x)\, \zeta(dx), \quad A \in \mathfrak{B}.$$

Alors

$$\mathsf{E}\lambda(A)\overline{\lambda(B)} = \int \chi_A(x)\chi_B(x)\,|g(x)|^2\, m(dx) = \int\limits_{A\cap B} |g(x)|^2\, m(dx).$$

Si l'on munit $\mathfrak{B}$ de la nouvelle mesure

$$l(A) = \int\limits_A |g(x)|^2\, m(dx),$$

on voit que $\lambda(A)$ sera une mesure stochastique orthogonale de fonction structurelle $l(A)$.

LEMME 1. *Si* $f(x) \in \mathcal{L}_2\{l\}$, *alors* $f(x)\,g(x) \in \mathcal{L}_2\{m\}$ *et*

$$\int f(x)\,\lambda(dx) = \int f(x)\,g(x)\,\zeta(dx).$$

Démonstration. Le lemme est évident pour les fonctions simples $f(x) = \sum\limits_k c_k \chi_{A_k}(x)$, $A_k \in \mathfrak{B}$. D'autre part, si $f_k(x)$ est une suite fondamentale de fonctions simples dans $\mathcal{L}_2\{l\}$, alors

$$\left|\int f_n(x)\,\lambda(dx) - \int f_{n+m}(x)\,\lambda(dx)\right|^2 = \int |f_n(x) - f_{n+m}(x)|^2\, l(dx) =$$
$$= \int |f_n(x) - f_{n+m}(x)|^2\, |g(x)|^2\, m(dx),$$

i.e. $f_n(x)\,g(x)$ est fondamentale dans $\mathcal{L}_2\{m\}$. En passant à la limite pour $n \to \infty$ dans

$$\int f_n(x)\,\lambda(dx) = \int f_n(x)\,g(x)\,\zeta(dx)$$

on obtient ce qu'on voulait dans le cas général. ■

LEMME 2. *Si* $A \in \mathfrak{B}_0$,

$$\zeta(A) = \int \frac{\chi_A(x)}{g(x)}\,\lambda(dx).$$

On remarquera tout d'abord que $g(x) = 0$ sur un ensemble de l-mesure 0; donc $\frac{1}{g(x)} \neq \infty$ (mod l). Par ailleurs

$$\int \frac{\chi_A(x)}{|g(x)|^2}\, l(dx) = \int \frac{1}{|g(x)|^2}\, |g(x)|^2\, m(dx) = m(A) < \infty.$$

On peut donc se servir du lemme 1 :

$$\int \frac{1}{g(x)} \chi_A(x) \lambda(dx) = \int \frac{1}{g(x)} \chi_A(x) g(x) \zeta(dx) = \zeta(A). \blacksquare$$

Soient T intervalle fini ou infini de la droite, $\mathfrak{B}^1$ tribu de parties de T mesurables-Lebesgue, l mesure de Lebesgue.

Supposons que $g(t, x)$ est $\mathfrak{B}^1 \times \mathfrak{B}$-mesurable, $g(t, x) \in \in \mathcal{L}_2\{l \times m\}$ et $g(t, x) \in \mathcal{L}_2\{m\}$ pour tout $t \in T$. Considérons l'intégrale stochastique

$$\xi(t) = \int g(t, x) \zeta(dx). \tag{10}$$

Elle est presque sûrement définie pour tout t.

LEMME 3. *L'intégrale stochastique* (10) *peut être définie comme une fonction de t de telle sorte que le processus* $\xi(t)$ *soit mesurable.*

Démonstration. Si

$$g(t, x) = \sum c_k \chi_{B_k}(t) \chi_{A_k}(x), \tag{11}$$

$B_k \in \mathfrak{B}^1$, $A_k \in \mathfrak{B}$, alors $\xi(t) = \sum c_k \chi_{B_k}(t) \zeta(A_k)$ est une fonction $\mathfrak{B}^1 \times \mathfrak{S}$-mesurable de (t, ω), $t \in T$, $\omega \in \Omega$. Dans le cas général on peut composer une suite de fonctions simples $g_n(t, x)$ de la forme (11) telles que

$$\int\int |g(t, x) - g_n(t, x)|^2 m(dx)\, dt \to 0 \quad \text{pour } n \to \infty.$$

Soit $\xi_n(t)$ une suite de processus construits à l'aide de la formule (10) pour $g = g_n$. Il existe alors un processus $\tilde{\xi}(t)$ tel que

$$\int \mathsf{E} |\tilde{\xi}(t) - \xi_n(t)|^2 dt \to 0 \quad \text{pour } n \to \infty$$

et $\tilde{\xi}(t)$ est une fonction $\mathfrak{B}^1 \times \mathfrak{S}$-mesurable de (t, ω). D'autre part,

$$\int \mathsf{E} |\xi(t) - \xi_n(t)|^2 dt = \int \int |g(t, x) - g_n(t, x)|^2 m(dx)\, dt \to 0,$$

d'où il suit que $\mathsf{E} |\xi(t) - \tilde{\xi}(t)|^2 = 0$ presque pour tous les t. Posons

$$\xi'(t) = \begin{cases} \tilde{\xi}(t) & \text{si } \mathsf{P}\{\xi(t) \neq \tilde{\xi}(t)\} = 0, \\ \xi(t) & \text{si } \mathsf{P}\{\xi(t) \neq \xi(t)\} > 0. \end{cases}$$

Le processus $\xi'(t)$ est mesurable (car $\xi'(t)$ coïncide avec la fonction $\mathfrak{B}^1 \times \mathfrak{S}$-mesurable $\tilde{\xi}(t)$ à un ensemble près de mesure nulle) et est stochastiquement équivalent à $\xi(t)$. $\blacksquare$

Dans la suite on supposera que les processus définis par des intégrales stochastiques (10) et vérifiant les conditions énumérées précédemment sont mesurables.

LEMME 4. *Si $g(t, s)$ et $h(t)$ sont des fonctions boréliennes,*

$$\int_a^b \int_{-\infty}^{\infty} |g(t, s)|^2 dtm(ds) < \infty, \int_a^b |h(t)|^2 dt < \infty, \qquad (12)$$

ζ une mesure stochastique orthogonale sur $\{\mathscr{R}^1, \mathfrak{B}^1\}$, alors

$$\int_a^b h(t) \int_{-\infty}^{\infty} g(t, s) \zeta(ds) dt = \int_{-\infty}^{\infty} g_1(s) \zeta(ds), \qquad (13)$$

où

$$g_1(s) = \int_a^b h(t) g(t, s) dt.$$

Démonstration. L'espérance mathématique du carré du module de l'intégrale du premier membre de (13) vaut:

$$\int_a^b \int_a^b h(t_1)\overline{h(t_2)} \left(\int_{-\infty}^{\infty} g(t_1, s)\overline{g(t_2, s)}\, m(ds)\right) dt_1 dt_2 =$$

$$= \int_{-\infty}^{\infty} \left| \int_a^b h(t) g(t, s) dt \right|^2 m(ds) \leqslant \int_a^b |h(t)|^2 dt \times$$

$$\times \int_{-\infty}^{\infty} \int_a^b |g(t, s)|^2 dtm(ds).$$

Pour l'espérance mathématique du carré du module du second membre de (13) on a l'inégalité

$$\int_{-\infty}^{\infty} \left| \int_a^b h(t) g(t, s) dt \right|^2 m(ds) \leqslant \int_a^b |h(t)|^2 dt \cdot \int_{-\infty}^{\infty} \int_a^b |g(t, s)|^2 dtm(ds).$$

Donc les deux membres de (13) sont continus par rapport au passage à la limite sur les suites $g_n(t, s)$ convergentes dans $\mathscr{L}_2\{\Phi\}$, où Φ est le produit cartésien de la mesure de Lebesgue par la mesure m dans la bande $[a, b] \times]-\infty, \infty[$. D'autre part, l'ensemble des fonctions $g(t, s)$ vérifiant (13) est un espace vectoriel et contient toutes les fonctions de la forme $\sum c_k \chi_{A_k}(t) \chi_{B_k}(s)$. Il contient donc toutes les fonctions de $\mathscr{L}_2\{\Phi\}$. ∎

REMARQUE. *Si les conditions du lemme 4 sont réunies pour tout intervalle fini $]a, b[$ et existe l'intégrale*

$$\int_{-\infty}^{\infty} h(t) g(t, s) dt = \lim_{\substack{a \to -\infty \\ b \to +\infty}} \int_a^b h(t) g(t, s) dt$$

au sens de la convergence dans $\mathscr{L}_2\{m\}$, *alors*

$$\int_{-\infty}^{\infty} h(t) \int_{-\infty}^{\infty} g(t, s) \zeta(ds)\, dt = \int_{-\infty}^{\infty} f_1(s) \zeta(ds), \tag{14}$$

où

$$f_1(s) = \int_{-\infty}^{\infty} h(t)\, g(t, s)\, dt.$$

Ce résultat découle immédiatement du fait que le premier membre de (14) est la limite en m.q. du premier membre de (13) et de la possibilité de passer à la limite sous le signe de l'intégrale stochastique dans le second membre de (13). ■

Généralisons les résultats précédents aux mesures stochastiques vectorielles. On se bornera au cas élémentaire de l'intégration de fonctions scalaires qui diffère peu de l'intégration par rapport à des mesures stochastiques numériques.

Soit Z un espace vectoriel complexe de dimension p. Pour simplifier on admettra qu'une base a été figée dans cet espace. Supposons qu'à tout $\Delta \in \mathfrak{M}$ est associée une variable aléatoire vectorielle $\zeta(\Delta)$ à valeurs dans Z^p, $\zeta(\Delta) = \{\zeta^1(\Delta), \zeta^2(\Delta), \ldots, \zeta^p(\Delta)\}$. Soit $|\zeta(\Delta)|$ la norme du vecteur $\zeta(\Delta)$

$$|\zeta(\Delta)|^2 = \sum_{k=1}^{p} |\zeta^k(\Delta)|^2.$$

Supposons que

1) $\mathsf{E}\,|\zeta(\Delta)|^2 < \infty$, $\zeta(\varnothing) = 0$;

2) $\zeta(\Delta_1 \cup \Delta_2) = \zeta(\Delta_1) + \zeta(\Delta_2)$ (mod P) si $\Delta_1 \cap \Delta_2 = \varnothing$;

3) $\mathsf{E}\zeta^k(\Delta_1)\overline{\zeta^j(\Delta_2)} = m_j^k(\Delta_1 \cap \Delta_2)$, $\Delta_i \in \mathfrak{M}$, $i = 1, 2$; $k, j = 1, \ldots, p$.

La famille de vecteurs aléatoires $\{\zeta(\Delta), \Delta \in \mathfrak{M}\}$ sera dite *mesure stochastique (orthogonale) vectorielle élémentaire* et la matrice $m(\Delta) = \{m_j^k(\Delta)\} = \mathsf{E}\zeta(\Delta)\zeta^*(\Delta)$ *matrice structurelle.*

On remarquera que, comme fonction de Δ_1 et Δ_2, la matrice $m(\Delta_1 \cap \Delta_2)$ possède les propriétés de la matrice de corrélation d'une fonction aléatoire vectorielle. De plus, si $\Delta_1 \cap \Delta_2 = \varnothing$, alors

$$m(\Delta_1 \cup \Delta_2) = m(\Delta_1) + m(\Delta_2).$$

D'où il suit que les éléments diagonaux de la matrice $m(\Delta)$ sont des prémesures. D'autre part, de l'inégalité

$$|m_j^k(\Delta)| \leqslant \sqrt{m_k^k(\Delta)\, m_j^j(\Delta)} \tag{15}$$

il suit

$$\sum_r |m_j^k(\Delta_r)| \leqslant \{\sum_r m_k^k(\Delta_r) \sum_r m_j^j(\Delta_r)\}^{1/2}, \tag{16}$$

et par suite les fonctions d'ensembles m_j^k $(k, j = 1, \ldots, p)$ sont à variation bornée sur Δ.

Posons $m_0(\Delta) = \operatorname{Tr} m(\Delta) = \sum_{k=1}^{p} m_k^k(\Delta)$. De (16) il suit que si $\sum_{r=1}^{m_N} m_0(\Delta_r^N) \to 0$ pour $N \to \infty$, alors $\sum_{r=1}^{m_N} |m_j^k(\Delta_r^N)| \to 0$. D'où il suit que les fonctions $m_j^k(\Delta)$ sont prolongeables à des fonctions dénombrablement additives d'ensembles sur $\mathfrak{B}$ si la fonction $m_0(\Delta)$ est semi-additive sur $\mathfrak{M}$.

Dans la suite les fonctions matricielles ainsi obtenues à partir de la fonction structurelle d'une mesure stochastique orthogonale élémentaire seront dites *mesures matricielles définies positives*.

Plus haut par $\mathfrak{B}$ on a désigné la complétion de $\sigma\{\mathfrak{M}\}$ par rapport à la mesure élémentaire prolongée $m_0(\Delta)$. Pour simplifier on conservera les mêmes notations pour les prolongements des fonctions m_j^k, m_0 et la matrice m sur $\mathfrak{B}$ et on admettra que $m_0(\Delta)$ est semi-additive sur $\mathfrak{M}$.

Définissons sur $\mathscr{L}_0\{\mathfrak{M}\}$ une intégrale stochastique par la formule

$$\eta = \int f(x)\,\zeta(dx) = \sum_{k=1}^{n} c_k \zeta(\Delta_k) \tag{17}$$

si $f(x) = \sum_{k=1}^{n} c_k \chi_{\Delta_k}(x)$, $\Delta_k \in \mathfrak{M}$ $(k = 1, \ldots, n)$. Les valeurs de cette intégrale forment un vecteur aléatoire (vecteur colonne) à valeurs dans Z^p. Soit $\mathscr{L}_0^p(\zeta)$ l'ensemble de tous les vecteurs aléatoires η de la forme (17). Si $g(x) = \sum_{k=1}^{n} d_k \chi_{\Delta_k}(x)$, alors

$$\mathsf{E}\left(\int f(x)\,\zeta(dx)\left(\int g(x)\,\zeta(dx)\right)^*\right) = \sum_{k=1}^{n} c_k \bar{d}_k m(\Delta_k),$$

qui s'écrit encore

$$\mathsf{E}\left(\int f(x)\,\zeta(dx)\left(\int g(x)\,\zeta(dx)\right)^*\right) = \int f(x)\,\overline{g(x)}\,m(dx). \tag{18}$$

Ce qui entraîne

$$\mathsf{E}\left|\int f(x)\,\zeta(dx)\right|^2=\int |f(x)|^2\,m_0(dx). \tag{19}$$

Munissons $\mathscr{L}_0\{\mathfrak{M}\}$ du produit scalaire

$$(f,\ g)=\int f(x)\,\overline{g(x)}\,m_0(dx).$$

La formule (17) définit une application isométrique $\eta=\psi(f)$ de l'espace $\mathscr{L}_0\{\mathfrak{M}\}$ sur $\mathscr{L}_0^p\{\zeta\}$ si dans $\mathscr{L}_0^p\{\zeta\}$ le produit scalaire des éléments η_1 et η_2 est défini comme $\mathsf{E}\eta_2^*\eta_1$. Appelons $\mathscr{L}_2^p\{\zeta\}$ la fermeture de l'espace des variables aléatoires $\mathscr{L}_0^p\{\zeta\}$ et $\mathscr{L}_2\{\mathfrak{M}\}$ la complétion de $\mathscr{L}_0\{\mathfrak{M}\}$.

Par analogie à l'inégalité (16) considérons l'inégalité

$$\int |f(x)|\,|m_j^k|(dx)\leqslant\left\{\int |f(x)|\,m_k^k(dx)\int |f(x)|\,m_j^j(dx)\right\}^{1/2}, \tag{20}$$

où $|m_j^k|(A)$ est la variation absolue de la fonction m_j^k. L'inégalité (20) entraîne l'existence et la continuité de l'intégrale

$$\int f(x)\,\overline{g(x)}\,m_j^k(dx)$$

comme fonctionnelle de f et g dans $\mathscr{L}_2\{m_0\}$

De là on déduit que l'application isométrique $\eta=\psi(f)$ de $\mathscr{L}_0\{\mathfrak{M}\}$ sur $\mathscr{L}_0^p\{\zeta\}$ est prolongeable à une application isométrique de $\mathscr{L}_2\{\mathfrak{M}\}$ sur $\mathscr{L}_2^p\{\zeta\}$. Cela étant, le vecteur aléatoire η est appelé intégrale stochastique et noté

$$\eta=\int f(x)\,\zeta(dx),$$

où $f(x)\in\mathscr{L}_2\{m_0\}$.

On définirait de façon analogue la mesure stochastique vectorielle $\tilde{\zeta}(A)$.

§ 3. Représentations intégrales de fonctions aléatoires

En se servant des résultats du paragraphe précédent on obtient diverses représentations des fonctions aléatoires par des intégrales stochastiques.

Supposons tout d'abord qu'une fonction vectorielle p-dimensionnelle $\xi(\theta)$, $\theta\in\Theta$, est représentable par

$$\xi(\theta)=\int g(\theta,\ x)\,\zeta(dx), \tag{1}$$

où ζ est une mesure stochastique sur l'espace probabilisable $\{X,\ \mathfrak{B}\}$ à valeurs dans Z^p et de matrice structurelle $m(A)$ (nous conservons les notations du paragraphe précédent), que $g(\theta,\ x)$ est une fonction

scalaire et pour tout $\theta \in \Theta$

$$g(\theta, x) \in \mathscr{L}_2\{m_0\} = \mathscr{L}_2\{X, \mathfrak{B}, m_0\}, \quad m_0(A) = \operatorname{Tr} m(A).$$

En vertu de la formule (18), § 2, la matrice de covariance de la fonction aléatoire $\xi(\theta)$ s'écrit

$$B(\theta_1, \theta_2) = \mathsf{E}\xi(\theta_1)\xi^*(\theta_2) = \int g(\theta_1, x)\overline{g(\theta_2, x)}\, m(dx), \tag{2}$$

et de (19), § 2, il suit que

$$\mathsf{E}\xi^*(\theta_2)\xi(\theta_1) = \int g(\theta_1, x)\overline{g(\theta_2, x)}\, m_0(dx). \tag{3}$$

On rappelle que $\{X, \mathfrak{B}, m_0\}$ est un espace muni d'une mesure complète, $\mathscr{L}_2\{m_0\}$ est un espace hilbertien de fonctions $\mathfrak{B}$-mesurables à valeurs complexes de carré m_0-intégrables.

Appelons $\mathscr{L}_2\{g\}$ la fermeture dans $\mathscr{L}_2\{m_0\}$ de l'enveloppe linéaire engendrée par le système de fonctions $\{g(\theta, x), \theta \in \Theta\}$. $\mathscr{L}_2\{g\}$ est alors un sous-espace vectoriel fermé de $\mathscr{L}_2\{m_0\}$. Si $\mathscr{L}_2\{g\} = \mathscr{L}_2\{m_0\}$, on dit que le système de fonctions $\{g(\theta, x), \theta \in \Theta\}$ est complet dans $\mathscr{L}_2\{m_0\}$.

Soient $\{\xi(\theta), \theta \in \Theta\}$ une fonction aléatoire hilbertienne à valeurs dans Z^p, $\mathscr{L}_0\{\xi\}$ l'ensemble des vecteurs aléatoires

$$\eta = \sum_{k=1}^{n} c_k \xi(\theta_k), \quad n = 1, 2, \ldots, \theta_k \in \Theta, \tag{4}$$

où c_k sont des nombres complexes quelconques, $\mathscr{L}_2\{\xi\}$ la fermeture de $\mathscr{L}_0\{\xi\}$ au sens de la convergence en m.q. des vecteurs aléatoires.

DÉFINITION. *Une famille de vecteurs aléatoires* $\{\eta_\alpha, \alpha \in A\}$, $\eta_\alpha \in \mathscr{L}_2\{\Omega, \mathfrak{S}, \mathsf{P}\}$ *est dite attachée à une fonction aléatoire* $\{\xi(\theta), \theta \in \Theta\}$ *si* $\eta_\alpha \in \mathscr{L}_2\{\xi\}$, $\alpha \in A$.

THÉORÈME 1. *Supposons que la matrice de covariance d'une fonction aléatoire* $\{\xi(\theta), \theta \in \Theta\}$ *admette la représentation* (2), *où* m *est une mesure matricielle définie positive sur* $\{X, \mathfrak{B}\}$, $g(\theta, x) \in \mathscr{L}_2\{m_0\}$, $\theta \in \Theta$, *et la famille* $\{g(\theta, x), \theta \in \Theta\}$ *est complète dans* $\mathscr{L}_2\{X, \mathfrak{B}, m_0\}$. *Alors* $\xi(\theta)$ *est représentable par la formule* (1), *où* $\{\zeta(B), B \in \mathfrak{B}\}$ *est une mesure vectorielle orthogonale stochastique attachée à la fonction aléatoire* $\xi(\theta)$ *de fonction structurelle* $m(\cdot)$, *et l'égalité* (1) *est réalisée presque sûrement pour tout* θ.

Démonstration. A toute combinaison linéaire

$$f(x) = \sum_{k=1}^{n} c_k g(\theta_k, x), \quad \theta_k \in \Theta, \tag{5}$$

associons un vecteur aléatoire η, $\eta = \psi(f)$ à l'aide de (4). Désignons par $\mathscr{L}_0\{g\}$ l'ensemble des fonctions (5). Munissons $\mathscr{L}_0\{g\}$ du produit scalaire

$$(f_1, f_2) = \int f_1(x)\overline{f_2(x)}\, m_0(dx). \tag{6}$$

L'application $\eta = \psi(f)$ est une isométrie de $\mathscr{L}_0\{g\}$ sur $\mathscr{L}_0\{\xi\}$. Elle peut être prolongée à une isométrie de $\mathscr{L}_2\{g\}$ sur $\mathscr{L}_2\{\xi\}$. Si $B \in \mathfrak{B}$, alors $\chi_B(x) \in \mathscr{L}_2\{m_0\} = \mathscr{L}_2\{g\}$ en vertu de la complétude de la famille $\{g(\theta, x), \theta \in \Theta\}$. Posons $\zeta(A) = \psi(\chi_A)$. Alors $\zeta(A)$ est une mesure stochastique vectorielle dont la fonction structurelle est confondue avec m:

$$\mathsf{E}\zeta(A_1)\zeta^*(A_2) = \int \chi_{A_1}(\theta)\overline{\chi_{A_2}(\theta)}\, m(d\theta) = m(A_1 \cap A_2).$$

Soit maintenant une fonction aléatoire $\tilde{\xi}(\theta)$ définie par l'intégrale stochastique

$$\tilde{\xi}(\theta) = \int g(\theta, x)\,\zeta(dx).$$

Comme

$$\mathsf{E}\xi(\theta)\zeta^*(A) = \int g(\theta, x)\chi_A(x)\, m(dx),$$

l'isométrie $\eta = \psi(f)$ implique:

$$\mathsf{E}\xi(\theta)\tilde{\xi}^*(\theta) = \int g(\theta, x)\overline{g(\theta, x)}\, m(dx).$$

D'où il vient

$$\mathsf{E}\,|\xi(\theta) - \tilde{\xi}(\theta)|^2 =$$

$$= \mathsf{E}\xi^*(\theta)\,\xi(\theta) - \mathsf{E}\tilde{\xi}^*(\theta)\,\xi(\theta) - \mathsf{E}\xi^*(\theta)\tilde{\xi}(\theta) + \mathsf{E}\tilde{\xi}^*(\theta)\tilde{\xi}(\theta) = 0,$$

d'où le théorème. ■

Voici quelques exemples d'application de ce théorème. Pour abréger on conviendra jusqu'à la fin de ce paragraphe de désigner simplement par stationnaires les processus stationnaires au sens large.

La matrice de corrélation d'un processus stationnaire et continu en m. q. est représentable par (cf. § 5, chapitre I)

$$R(t_1, t_2) = R(t_1 - t_2) = \int_{-\infty}^{\infty} e^{iu(t_1 - t_2)} F(du), \tag{7}$$

où $F(\cdot)$ est une mesure matricielle définie non négative (matrice spectrale du processus). La formule (7) est un cas particulier de la formule (2) dans laquelle les fonctions $g(\theta, x)$ correspondent à e^{iut},

$\theta \leftrightarrow t$, $x \to u$; de plus, l'ensemble de fonctions $\{e^{iut}, -\infty < u < \infty\}$ est complet dans $\mathscr{L}_2\ \{m_0\}$, où m_0 est une mesure quelconque bornée sur la droite. On est donc dans les conditions du théorème 1 et l'on obtient le

Théorème 2. *Tout processus aléatoire vectoriel stationnaire continu en m. q.* $\xi(t)$ $(-\infty < t < \infty)$, $\mathsf{E}\xi(t) = 0$, *est représentable par*

$$\xi(t) = \int_{-\infty}^{\infty} e^{iut}\zeta(du), \tag{8}$$

où $\zeta(A)$ *est une mesure stochastique orthogonale vectorielle sur* $\mathfrak{B}$ *attachée à* $\xi(t)$. *Entre* $\mathscr{L}_2\ \{\xi\}$ *et* $\mathscr{L}_2\ \{F_0\}$, *où* $F_0(\cdot) = \operatorname{Tr} F(\cdot)$ *il existe une isométrie telle que*

a) $\xi(t) \leftrightarrow e^{iut}$, $\zeta(A) \leftrightarrow \chi_A(u)$;

b) *si* $\eta_i \leftrightarrow g_i(u)$ $(i = 1, 2)$, *alors*

$$\eta_i = \int g_i(u)\,\zeta(du)$$

et

$$\mathsf{E}\eta_1\eta_2^* = \int g_1(u)\overline{g_2(u)}\,F(u).$$

La formule (8) s'appelle *représentation spectrale* du processus stationnaire, la mesure $\zeta(A)$ *mesure spectrale stochastique* du processus. Le théorème 2 nous dit que

$$\mathsf{E}\zeta(A_1)\zeta^*(A_2) = \int_{A_1\cap A_2} F(du) = F(A_1 \cap A_2). \tag{9}$$

i. e. $F(\cdot)$ est fonction structurelle de la mesure stochastique vectorielle $\zeta(\cdot)$.

Remarque 1. Pour tout $\eta \in \mathscr{L}_2\ \{\xi\}$ on a $\mathsf{E}\eta = 0$. En particulier, pour tout $A \in \mathfrak{B}$, $\mathsf{E}\zeta(A) = 0$.

Remarque 2. Si $\mathsf{E}\xi(t) = a \neq 0$, le processus $\xi(t) - a$ est justiciable du théorème précédent. D'autre part, la représentation (8) reste valable dans le cas général pourvu que l'on ajoute à $\zeta(A)$ une mesure concentrée au point $u = 0$ et égale à a.

A titre d'applications du théorème 2 on se propose d'établir la formule de Kotelnikov-Shannon pour un processus aléatoire unidimensionnel dont la mesure spectrale est concentrée sur un intervalle fini $[-B, B]$. Développons la fonction e^{iut} en série de Fourier sur l'intervalle $[-B, B]$. On obtient

$$e^{iut} = \sum_{n=-\infty}^{\infty} \frac{\sin(Bt - \pi n)}{Bt - n\pi}\, e^{i\frac{\pi n}{B}u}.$$

Etant uniformément convergente en u sur tout intervalle $[-B', B']$, $B' < B$, et possédant des sommes partielles, cette série est convergente dans $\mathscr{L}_2\{m_0\}$. L'isomorphisme des espaces $\mathscr{L}_2\{m_0\}$ et $\mathscr{L}_2\{\xi\}$ implique (au sens de la convergence en m. q.)

$$\xi(t) = \sum_{n=-\infty}^{\infty} \frac{\sin(Bt - \pi n)}{Bt - \pi n} \xi\left(\frac{\pi n}{B}\right). \tag{10}$$

Donc à tout instant t la valeur de la fonction aléatoire $\xi(t)$ est donnée par ses valeurs aux instants équidistants $\pi n/B$, $n = 0, \pm 1, \pm 2, \ldots$

On pourrait énoncer un théorème complètement analogue au théorème 2 pour les suites vectorielles stationnaires ξ_n, $n = 0, \pm 1, \pm 2, \ldots$ La seule différence cependant est que la mesure spectrale de la suite est concentrée sur l'intervalle $[-\pi, \pi[$ et non sur la droite réelle tout entière comme dans le cas d'un processus à temps continu (cf. théorème 1, § 2).

Des théorèmes 1 et 2 du § 2 on déduit la généralisation suivante du théorème 2 de la représentation spectrale d'un champ homogène continu en m. q.

THÉORÈME 3. *Tout champ vectoriel homogène continu en m. q. $\xi(x)$, $x \in \mathscr{R}^m$, est représentable par*

$$\xi(x) = a + \int_{\mathscr{R}^m} e^{i(x, u)} \zeta(du), \quad a = \mathsf{E}\xi(x),$$

où ζ est une mesure vectorielle orthogonale sur $\mathfrak{B}^m$ attachée au champ $\xi(x)$. Entre $\mathscr{L}_2\{\xi\}$ et $\mathscr{L}_2\{F_0\}$, $F_0(\cdot) = \operatorname{Tr} F(\cdot)$ il existe une isométrie telle que

a) $\zeta(x) \leftrightarrow e^{i(x, u)}$;

b) *si* $\eta_i \leftrightarrow g_i(u)$, $\eta_i \in \mathscr{L}_2\{\xi\}$, $g_i(u) \in \mathscr{L}_2\{F_0\}$, $i = 1, 2$, *alors*

$$\eta_i = \int_{\mathscr{R}^m} g_i(u) \zeta(du),$$

$$\mathsf{E}\eta_1\eta_2^* = \int_{\mathscr{R}^m} g_1(u) \overline{g_2(u)} F(du).$$

COROLLAIRE. *Si un champ homogène (scalaire) $\xi(x)$, $\mathsf{E}\xi(x) = 0$, possède un spectre borné, c'est-à-dire*

$$R(x) = \int_{-B_1}^{B_1} \cdots \int_{-B_m}^{B_m} e^{i(x, u)} F(du),$$

alors il est défini de façon unique par ses valeurs aux points du réseau

$$\left\{x_n = \frac{\pi n^1}{B_1}, \frac{\pi n^2}{B_2}, \ldots, \frac{\pi n^m}{B_m}, \; n^k = 0, \pm 1, \pm 2, \ldots\right\}$$

par la formule

$$\xi(x) = \sum_{n=(n^1, \ldots, n^m)} \prod_{k=1}^{m} \frac{\sin(B_k x^k - \pi n^k)}{B_k x^k - \pi n^k} \xi\left(\frac{\pi n^1}{B_1}, \frac{\pi n^2}{B_2}, \ldots, \frac{\pi n^m}{B_m}\right), \quad (11)$$

dans laquelle la sommation est étendue à tous les vecteurs entiers n et la série du second membre est convergente en m. q. pour tout x.

Considérons encore la représentation spectrale d'un champ aléatoire isotrope bidimensionnel continu en m.q. La formule (10) du § 5, chapitre I, montre que la fonction de corrélation du champ est de la forme

$$R(x_1, x_2) = R(\rho) = \int_0^\infty J_0(u\rho)\, g(du), \quad (12)$$

où x_1 et x_2 sont des points du plan, ρ leur distance. Si (r_i, θ_i) sont les coordonnées polaires du point x_i $(i = 1, 2)$, alors

$$\rho = \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos(\theta_1 - \theta_2)}.$$

La formule de sommation pour la fonction J_0:

$$J_0(u\rho) = \sum_{k=-\infty}^{\infty} J_k(ur_1)\, J_k(ur_2)\, e^{ik(\theta_1 - \theta_2)}$$

nous permet d'écrire (12) sous la forme

$$R(\rho) = \int_0^\infty \int_{-\infty}^\infty J_\nu(ur_1)\, e^{i\nu\theta_1} J_\nu(ur_2)\, e^{i\nu\theta_2} g(du)\, \varepsilon(d\nu),$$

où $\varepsilon(d\nu)$ est une mesure concentrée aux points $k = 0, \pm 1, \pm 2, \ldots$ et de plus $\varepsilon(\{k\}) = 1$. Le théorème 1 nous apprend que le champ homogène isotrope plan continu en m.q. $\xi(x)$, $x = re^{i\theta}$ $(\mathsf{E}\xi(x) = 0)$ admet la représentation

$$\xi(x) = \sum_{k=-\infty}^{\infty} e^{ik\theta} \int_0^\infty J_k(ur)\, \zeta_k(du), \quad (13)$$

où ζ_k est une suite de mesures stochastiques orthogonales sur la section $[0, \infty[$.

§ 4. Transformations linéaires

Imaginons-nous un système Σ (un appareil ou un dispositif) destiné à transformer des signaux (fonctions) $x(t)$ dépendant du temps t. La fonction à transformer est dite *fonction d'entrée du système*; la fonction transformée, *fonction de sortie* ou *réponse* à la fonction

d'entrée. Mathématiquement tout système est donné par une classe $\mathscr{D}$ de fonctions « admissibles » d'entrée et une relation de la forme

$$z(t) = T(x \mid t),$$

où $x = x(s)$ $(-\infty < s < \infty)$ est une fonction d'entrée, $x(s) \in \mathscr{D}$, $z(t)$ la valeur de la fonction de sortie à l'instant t.

Le système Σ est linéaire si:

a) la classe de fonctions admissibles $\mathscr{D}$ est linéaire;

b) l'opérateur T vérifie le principe de superposition

$$T(\alpha x_1 + \beta x_2 \mid t) = \alpha T(x_1 \mid t) + \beta T(x_2 \mid t).$$

Considérons la translation dans le temps S_τ $(-\infty < \tau < \infty)$ définie par

$$x_\tau(t) = S_\tau(x \mid t) = x(t + \tau).$$

Elle est définie sur l'ensemble des fonctions de $t \in]-\infty, \infty[$. Le système Σ est homogène si la classe de fonctions admissibles $\mathscr{D}$ est invariante par la translation S_τ: $S_\tau \mathscr{D} = \mathscr{D}$ et

$$T(x_\tau \| t) = T(x \mid t + \tau) \text{ ou } T(S_\tau x \mid t) = S_\tau T(x \mid t),$$

i.e. si la transformation T est permutable avec la translation S_τ $(-\infty < \tau < \infty)$.

Un exemple élémentaire de transformation linéaire est la transformation

$$z(t) = \int_{-\infty}^{\infty} h(t, s)\, x(s)\, ds, \tag{1}$$

pour laquelle la classe de fonctions admissibles $\mathscr{D}$ dépend des propriétés de $h(t, s)$. Supposons qu'à l'entrée du système est appliquée une fonction δ_{x-s} où δ_s est la fonction de Dirac. Alors $z(t) = h(t, s)$. Donc la fonction $h(t, s)$ doit être interprétée comme la réponse du système à la δ-fonction à l'instant s. Par suite $h(t, s)$ est dite *fonction de transfert impulsionnelle du système*. Si le système Σ est homogène, on a formellement

$$h(t, a - c) = T(\delta_{a-c} \mid t) = T(S_c\delta_a \mid t) = S_c T(\delta_a \mid t) = h(t + c, a)$$

ou, en remplaçant a par c et t par $t - c$,

$$h(t - c, 0) = h(t, c).$$

La fonction $h(t) = h(t + c, c)$ est dite *fonction de transfert impulsionnelle du système homogène*.

Dans le cas d'un système homogène, l'équation (1) s'écrit donc

$$z(t) = \int_{-\infty}^{\infty} h(t - s)\, x(s)\, ds. \tag{2}$$

L'expression (2) est appelée *produit de convolution* des fonctions $h(t)$ et $x(t)$.

Si la fonction de sortie diffère de la fonction d'entrée d'un facteur multiplicatif scalaire (la transformation T ne modifie pas la forme du signal),

$$T(f \mid t) = \lambda f(t) \quad (-\infty < t < \infty),$$

on dit que $f(t)$ est *fonction propre* et λ *valeur propre* de la transformation T. Les fonctions e^{iut} (u est un réel quelconque) sont propres pour les systèmes homogènes à fonction de transfert impulsionnelle intégrable. En effet, toutes les fonctions mesurables bornées sont admissibles et

$$\int_{-\infty}^{\infty} h(t-s)\, e^{ius}\, ds = \int_{-\infty}^{\infty} h(s)\, e^{iu(t-s)}\, ds = H(iu)\, e^{iut},$$

où

$$H(iu) = \int_{-\infty}^{\infty} h(s)\, e^{-isu}\, ds \tag{3}$$

la transformée de Fourier de la fonction de transfert impulsionnelle, est valeur propre de la transformation.

Donc le quotient de la réponse du système à la fonction harmonique simple e^{iut} par cette fonction, soit

$$H(iu) = \frac{T(e^{isu} \mid t)}{e^{iut}}$$

ne dépend pas du temps. La fonction $H(iu)$ est appelée *caractéristique fréquentielle* du système ou *coefficient de transfert*.

On peut donner une autre interprétation à la caractéristique fréquentielle du système (2) en considérant une autre classe de fonctions admissibles. Soit $x(t)$ intégrable. Le théorème de Fubini donne

$$\int_{-\infty}^{\infty} |z(t)|\, dt \leqslant \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |h(t-s)|\, |x(s)|\, ds\, dt =$$

$$= \int_{-\infty}^{\infty} x(s)|\, ds \int_{-\infty}^{\infty} |h(t)|\, dt < \infty,$$

i. e. la fonction $z(t)$ est également intégrable. Considérons la transformée de Fourier de la fonction $z(t)$. Le théorème de Fubini nous

dit que

$$\bar{z}(u)=\int_{-\infty}^{\infty} e^{-iut}z(t)\,dt=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-iu(t-s)}h(t-s)\,e^{-ius}x(s)\,ds\,dt=$$

$$=H(iu)\,\tilde{x}(u),\ \text{où}\ \tilde{x}(u)=\int_{-\infty}^{\infty} e^{ius}x(s)\,ds.$$

Donc le quotient de la transformée de Fourier de la fonction de sortie par la transformée de Fourier de la fonction d'entrée ne dépend pas de la fonction d'entrée du système et est égal à la caractéristique fréquentielle :

$$H(iu)=\frac{\tilde{z}(u)}{\tilde{x}(u)}.$$

Dans la formule (1) la réponse du système à l'instant t dépend des valeurs prises par la fonction d'entrée aussi bien aux instants $s<t$ qu'aux instants $s>t$. Dans les appareils il est impossible d'anticiper l'avenir. Aussi a-t-on

$$h(t,\ s)=0 \text{ pour } t<s. \tag{4}$$

La relation (4) est appelée *condition de réalisabilité physique*. Pour les systèmes vérifiant la condition (4) elle s'écrit

$$z(t)=\int_{-\infty}^{t} h(t,\ s)\,x(s)\,ds, \tag{5}$$

et dans le cas de systèmes homogènes

$$z(t)=\int_{-\infty}^{t} h(t-s)\,x(s)\,ds=\int_{0}^{\infty} h(s)\,x(t-s)\,ds. \tag{6}$$

Si la fonction est appliquée à l'entrée du système à partir de l'instant 0 ($x(s)=0$ pour $s<0$), on a

$$z(t)=\int_{0}^{t} h(t-s)\,x(s)\,ds. \tag{7}$$

La transformation de Laplace

$$\tilde{z}(p)=\int_{0}^{\infty} e^{-pt}z(t)\,dt \tag{8}$$

se prête mieux que celle de Fourier à l'étude de tels systèmes. De la formule (7) il suit

$$\tilde{z}(p)=H(p)\,\tilde{x}(p),\quad \tilde{x}(p)=\int_{0}^{\infty} e^{-pt}x(t)\,dt \tag{9}$$

pour Re $p \geqslant \alpha$ sous réserve que les fonctions $e^{-\alpha t}h(t)$ et $e^{-\alpha t}x(t)$ soient absolument intégrables.

Passons maintenant à l'objet principal de ce paragraphe: les transformations linéaires de processus aléatoires. On étudie essentiellement des transformations homogènes de processus stationnaires. Pour les cas plus généraux on se contentera de remarques simples.

Soit $\xi(t)$ un processus hilbertien mesurable $(-\infty < t < \infty)$ à covariance $B(t, s)$, la fonction $B(t, t)$ étant intégrable par rapport à t sur tout intervalle fini au même titre que la fonction $|h(s, t)|^2$ pour s fixe. Alors l'intégrale

$$\zeta(t) = \int_a^b h(t, s)\,\xi(s)\,ds \tag{10}$$

existe presque sûrement quels que soient a et b. Définissons l'intégrale impropre de $-\infty$ à ∞ comme la limite en m.q. des intégrales sur des intervalles finis:

$$\int_{-\infty}^{\infty} h(t, s)\,\xi(s)\,ds = \underset{\substack{a\to-\infty \\ b\to+\infty}}{\mathrm{l.i.m.}} \int_a^b h(t, s)\,\xi(s)\,ds.$$

Pour que cette limite existe il faut et il suffit que l'intégrale

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(t, s_1)\,B(s_1, s_2)\,\overline{h(t, s_2)}\,ds_1\,ds_2$$

existe comme intégrale impropre de Cauchy sur le plan. Si elle existe pour $t \in T$, alors $\zeta(t)$ est un processus aléatoire hilbertien sur T de covariance

$$B(t_1, t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(t_1, s_1)\,B(s_1, s_2)\,\overline{h(t_2, s_2)}\,ds_1\,ds_2.$$

Supposons maintenant que $\xi(t)$ est un processus stationnaire au sens large de mesure spectrale $F(du)$ et $\mathsf{E}\xi(t) = 0$. Cette hypothèse sera retenue jusqu'à la fin du présent paragraphe. L'intégrale

$$\eta(t) = \int_{-\infty}^{\infty} h(t-s)\,\xi(s)\,ds \tag{11}$$

existe (au sens défini précédemment) si et seulement si existe l'intégrale

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(t-s_1)R(s_1-s_2)\,\overline{h(t-s_2)}\,ds_1\,ds_2 =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(s_1)\,R(s_2-s_1)\,\overline{h(s_2)}\,ds_1\,ds_2,$$

où $R(t)$ est la fonction de corrélation du processus. Pour que ceci ait lieu il suffit maintenant que la fonction $h(t)$ soit absolument intégrable sur $]-\infty, \infty[$. En utilisant alors la représentation spectrale de la fonction de corrélation $R(t)$ on obtient l'expression suivante pour la fonction de corrélation $R_\eta(t_1, t_2)$ du processus $\eta(t)$:

$$R_\eta(t_1, t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(t_1 - s_1) R(s_1 - s_2) \overline{h(t_2 - s_2)}\, ds_1\, ds_2 =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(t_1 - s_1) e^{iu(s_1 - s_2)} \overline{h(t_2 - s_2)}\, ds_1\, ds_2 F(du) =$$

$$= \int_{-\infty}^{\infty} e^{i(t_1 - t_2)u} |H(iu)|^2 F(du) = R_\eta(t_1 - t_2).$$

Donc le processus $\eta(t)$ est aussi un processus stationnaire au sens large.

DÉFINITION. *Etant donné un processus* $\xi(t)$, *la transformation* T *est dite filtre admissible (ou simplement filtre) si elle est définie par la formule* (11), *où* $h(t)$ *est une fonction absolument intégrable sur* $]-\infty, \infty[$ *et de carré intégrable sur tout intervalle fini ou bien est limite en m.q. d'une suite de telles transformations (dans* $\mathcal{L}_2\{\xi\}$*).*

La condition de convergence de la suite de transformations (11) $\eta_n(t) = T_n(\xi \mid t)$ de fonctions de transfert impulsionnelles $h_n(t)$ et de caractéristiques fréquentielles $H_n(iu)$ s'écrit:

$$\mathsf{E}|\eta_n(t) - \eta_m(t)|^2 = \int_{-\infty}^{\infty} |H_n(iu) - H_m(iu)|^2 F(du) \to 0 \quad (12)$$

pour $n, m \to \infty$.

Ce qui signifie que la suite $H_n(iu)$ est fondamentale dans $\mathcal{L}_2\{F\}$. Ceci entraîne l'existence de $H(iu) = \text{l.i.m. } H_n(iu)$ (dans $\mathcal{L}_2\{F\}$) qui est appelée *caractéristique fréquentielle du filtre limite*, et si $\eta(t) = \text{l.i.m. } \eta_n(t)$, alors

$$R_\eta(t) = \int_{-\infty}^{\infty} e^{iut} |H(iu)|^2 F(du). \quad (13)$$

Inversement, toute fonction $H(iu) \in \mathcal{L}_2\{F\}$ est approchable au sens de la convergence dans $\mathcal{L}_2\{F\}$ par des fonctions qui sont les transformées de Fourier de fonctions absolument intégrables. Donc il est préférable de définir les filtres par leurs caractéristiques fréquentielles.

THÉORÈME 1. *Pour qu'une fonction* $H(iu)$ *soit caractéristique fréquentielle d'un filtre pour un processus* $\xi(t)$ *de mesure spectrale* F,

il faut et il suffit que $H(iu) \in \mathscr{L}_2\{F\}$. La fonction de corrélation du processus à la sortie d'un filtre de caractéristique fréquentielle $H(iu)$ est donnée par la formule (13).

Dans l'optique de l'interprétation énergétique de la fonction spectrale, de la formule (13) il suit que $|H(iu)|^2$ indique de combien de fois augmente l'énergie des composantes harmoniques simples à fréquences contenues dans l'intervalle $]u, u+du[$ en traversant le filtre.

THÉORÈME 2. *Si à l'entrée d'un filtre de caractéristique fréquentielle $H(iu)$ un processus $\xi(t)$ admet la représentation spectrale*

$$\xi(t) = \int_{-\infty}^{\infty} e^{iut}\zeta(du), \tag{14}$$

le processus $\eta(t)$ délivré à la sortie du filtre est de la forme

$$\eta(t) = \int_{-\infty}^{\infty} e^{iut} H(iu)\,\zeta(du). \tag{15}$$

En effet, si le filtre possède une fonction de transfert impulsionnelle absolument intégrable, alors

$$\eta(t) = \int_{-\infty}^{\infty} h(t-s)\,\xi(s)\,ds = \int_{-\infty}^{\infty} e^{iut}H(iu)\,\zeta(du).$$

Dans le cas général le théorème se démontre par passage à la limite sur les suites $H_n(iu)$ convergentes vers $H(iu)$ dans $\mathscr{L}_2\{F\}$. ■

Soit $\eta_k(t)$ un processus à la sortie d'un filtre de caractéristique fréquentielle $H_k(iu)$, $\mathsf{E}\eta_k(t) = 0$ $(k = 1, 2)$. Trouvons la fonction de corrélation des processus $\eta_1(t)$ et $\eta_2(t)$. L'isomorphisme des espaces $\mathscr{L}_2\{\zeta\}$ et $\mathscr{L}_2\{F\}$ implique immédiatement

$$R_{12}(t) = \mathsf{E}\eta_1(t+s)\overline{\eta_2(s)} = \int_{-\infty}^{\infty} e^{iut}H_1(iu)\overline{H_2(iu)}\,F(du). \tag{16}$$

Voici quelques exemples de filtres avec leurs caractéristiques fréquentielles.

1. Le *filtre à bandes* ne laisse passer (sans les modifier) que les composantes harmoniques des processus de fréquences u telles que $a < |u| < b$, $a > 0$. La caractéristique fréquentielle du filtre vaut $H(iu) = \chi_{(a,b)}(u) + \chi_{(-b,-a)}(u)$ et le filtre est admissible pour tout processus. La fonction de transfert impulsionnelle est donnée par la formule de Fourier:

$$h(t) = \frac{1}{2\pi}\left(\int_{-b}^{-a} + \int_{a}^{b}\right) e^{iut}\,du = \frac{\sin bt - \sin at}{\pi t}.$$

2. Le *filtre passe-hauts* supprime les fréquences basses. Sa caractéristique fréquentielle est $H(iu) = \chi_{\{|u|>a\}}(u)$. Il ne possède pas de fonction de transfert impulsionnelle.

3. Etudions la *dérivation en m.q. d'un processus stationnaire au sens large*. Pour que la dérivée en m.q. d'un processus $\xi(t)$ existe il suffit qu'existe $R''(0)$ (§ 3, corollaire 1). Il revient au même d'exiger (théorème 4, § 5, chapitre I) que

$$\int_{-\infty}^{\infty} u^2 F(du) < \infty. \tag{17}$$

D'autre part, si cette condition est réalisée, alors pour $h \to 0$

$$\frac{e^{iuh}-1}{h} \to iu \text{ (dans } \mathscr{L}_2\{F\})$$

et dans l'expression

$$\frac{\xi(t+h)-\xi(t)}{h} = \int_{-\infty}^{\infty} e^{iut}\,\frac{e^{iuh}-1}{h}\,\zeta(du)$$

on peut passer à la limite pour $h \to 0$ sous le signe de l'intégrale stochastique. Donc

$$\xi'(t) = \int_{-\infty}^{\infty} e^{iut} iu\zeta(du). \tag{18}$$

Par suite, la dérivation est réalisée par un filtre de caractéristique fréquentielle iu, admissible pour les processus stationnaires vérifiant (17). La fonction de transfert impulsionnelle n'existe pas, mais le filtre peut être considéré comme la limite ($\varepsilon \to 0$) de filtres à fonctions de transfert impulsionnelles $h_\varepsilon(t) = 0$ pour $|t| \geqslant \varepsilon$ et $h_\varepsilon(t) = -\frac{\operatorname{sgn} t}{\varepsilon^2}$ pour $|t| < \varepsilon$ auxquelles correspondent les caractéristiques fréquentielles $-\dfrac{4\sin^2\frac{u\varepsilon}{2}}{in\varepsilon^2}$.

4. *La translation dans le temps.* Comme

$$\xi(t+s) = \int_{-\infty}^{\infty} e^{iut} e^{ius}\zeta(du),$$

à la translation T_s, $T_s(\xi \mid t) = \xi(t+s)$, est associée la caractéristique fréquentielle $H(iu) = e^{ius}$. La fonction de transfert impulsionnelle n'existe pas.

5. *Equations différentielles.* Soit un filtre défini par une équation différentielle linéaire à coefficients constants

$$L\eta = M\xi, \tag{19}$$

où

$$L = a_0 \frac{d^n}{dt^n} + a_1 \frac{d^{n-1}}{dt^{n-1}} + \ldots + a_n,$$

$$M = b_0 \frac{d^m}{dt^m} + b_1 \frac{d^{m-1}}{dt^{m-1}} + \ldots + b_m.$$

L'équation (19) a un sens si seulement le processus $\xi(t)$ est m fois dérivable en m.q. On cherchera un processus stationnaire $\eta(t)$ n fois dérivable en m.q. vérifiant (19). On suppose que (19) admet une solution stationnaire. On peut décrire

$$\eta(t) = \int_{-\infty}^{\infty} e^{iut} H(iu)\, \zeta(du).$$

En appliquant les opérations M et L respectivement aux processus $\xi(t)$ et $\eta(t)$ on trouve

$$\int_{-\infty}^{\infty} e^{iut} L(iu)\, H(iu)\, \zeta(du) = \int_{-\infty}^{\infty} e^{iut} M(iu)\, \zeta(du),$$

où

$$L(iu) = \sum_{k=0}^{n} a_k (iu)^{n-k}, \qquad M(iu) = \sum_{k=0}^{\infty} b_k (iu)^{m-k},$$

d'où, si $L(iu)$ ne possède pas de racines réelles,

$$H(iu) = \frac{M(iu)}{L(iu)}. \tag{20}$$

Inversement, si le processus $\xi(t)$ est m fois dérivable en m.q., $M(iu) \in \mathscr{L}_2\{F\}$, $L(iu) \neq 0$ $(-\infty < u < \infty)$, alors le processus

$$\eta(t) = \int_{-\infty}^{\infty} e^{iut} \frac{M(iu)}{L(iu)}\, \xi(du)$$

est n fois dérivable en m.q. et est solution de l'équation (19). Donc, sous réserve que $M(iu) \in \mathscr{L}_2\{F\}$, $L(iu) \neq 0$, il existe un filtre unique correspondant à l'équation différentielle (19). Signalons toutefois que l'on peut résoudre l'équation (19) dans des cas plus généraux. Supposons que le polynôme $L(ui)$ ne possède pas de racines réelles. Il existe un filtre de caractéristique fréquentielle $M(iu)/L(iu)$ sans la condition $M(iu) \in \mathscr{L}_2\{F\}$; il suffit simplement d'exiger que $\frac{M(iu)}{L(iu)} \in \mathscr{L}_2\{F\}$, ce qui a toujours lieu pour $n \geqslant m$, où n est le degré du polynôme L. Donc pour $n \geqslant m$ un filtre de caractéristique fréquentielle (20), de dénominateur non nul pour aucun réel u, est admissible pour un processus d'entrée quelconque,

le processus de sortie étant identifié à une solution stationnaire de l'équation (19).

Limitons-nous comme précédemment aux équations différentielles pour lesquelles le polynôme $L(x)$ ne possède pas de racines imaginaires pures. Ecrivons le quotient $P(x)$ (non nul si $m \geqslant n$) de la fraction rationnelle $\frac{M(x)}{L(x)}$ et décomposons le reste en fractions élémentaires. On obtient

$$\frac{M(iu)}{L(iu)} = P(iu) + \sum_{k=1}^{n'} \sum_{s=1}^{l'_k} \frac{c'_{ks}}{(iu-p'_k)^s} + \sum_{k=1}^{n''} \sum_{s=1}^{l''_k} \frac{c''_{ks}}{(iu-p''_k)^s},$$

où $P(iu) = \sum_{k=0}^{m-n} d_k (iu)^k$ $(m \geqslant n)$ et $P(iu) = 0$ $(m < n)$, $\operatorname{Re} p'_k < 0$ et $\operatorname{Re} p''_k > 0$, p'_k et p''_k sont les zéros du polynôme $L(x) = 0$. Comme

$$\frac{1}{(iu-p)^s} = \frac{1}{(s-1)!} \frac{d^{s-1}}{dp^{s-1}} \int_0^\infty e^{pt} e^{-iut}\, dt = \int_0^\infty \frac{t^{s-1}}{(s-1)!} e^{pt} e^{-iut}\, dt$$

$$(\operatorname{Re} p < 0)$$

et

$$\frac{1}{(iu-p)^s} = -\int_{-\infty}^0 \frac{t^{s-1}}{(s-1)!} e^{pt} e^{-iut}\, dt \quad (\operatorname{Re} p > 0).$$

le processus $\eta(t)$ à la sortie du filtre s'écrit

$$\eta(t) = \sum_{k=0}^{m-n} d_k \xi^{(k)}(t) + \int_0^\infty \xi(t-\tau) G_1(\tau)\, d\tau + \int_0^\infty \xi(t+\tau) G_2(-\tau)\, d\tau,$$

où

$$G_1(t) = \sum_{k=1}^{n'} \left(\sum_{s=1}^{l'_k} \frac{c'_{ks} t^s}{(s-1)!} \right) e^{p'_k t} \qquad (t > 0)$$

$$G_2(t) = -\sum_{k=1}^{n''} \left(\sum_{s=1}^{l''_k} \frac{c''_{ks} t^s}{(s-1)!} \right) e^{p''_k t} \quad (t < 0).$$

On remarquera que si le polynôme $L(x)$ possède des racines à parties positives, le filtre correspondant est physiquement irréalisable.

§ 5. Filtres physiquement réalisables

Dans ce paragraphe on étudiera le problème suivant: quelles sont les fonctions spectrales délivrées à la sortie d'un filtre physiquement réalisable? On suppose qu'à l'entrée du filtre est appliqué un processus aléatoire élémentaire dans un certain sens.

Les processus étudiés dans ce paragraphe sont des processus unidimensionnels stationnaires au sens large. On omettra donc de le spécifier.

Commençons par les suites stationnaires. Nous n'étendrons pas aux suites les définitions et généralités euristiques relatives aux processus à temps continu, mais nous nous servirons tout de même de la terminologie appropriée. Imaginons-nous un système dont l'état à l'entrée et à la sortie est enregistré seulement à des instants entiers $t = 0, \pm 1, \pm 2, \ldots$

Supposons qu'à l'entrée du système est appliqué un signal unitaire à l'instant $t = 0$. Soit a_t la réponse du système à ce signal à l'instant t. Si le système n'anticipe pas l'avenir, alors $a_t = 0$ pour $t < 0$. Si le système est homogène en t, sa réponse au signal unitaire appliqué à l'instant s vaut a_{t-s}. La réponse d'un système linéaire homogène et physiquement réalisable à l'instant t à une suite de signaux ξ_n $(-\infty < n < \infty)$ sera

$$\eta(t) = \sum_{n=-\infty}^{t} a_{t-n}\xi(n) = \sum_{n=0}^{\infty} a_n\xi(t-n). \tag{1}$$

Ce processus est dit de *sommation glissante*.

Soit $\xi(n)$ une suite non corrélée de variables aléatoires et

$$\mathsf{E}\xi(n) = 0, \quad \mathsf{E}\xi(n)\overline{\xi(m)} = \delta_{nm} \quad (-\infty < n, \ m < \infty).$$

On l'appellera *suite standard*. Sa densité spectrale est constante.

Pour que la série (1) soit convergente en m.q. il faut et il suffit que

$$\sum_{n=0}^{\infty} |a_n|^2 < \infty. \tag{2}$$

Si cette condition est réalisée, le processus $\eta(t)$ est aussi stationnaire et

$$\mathsf{E}\eta(t) = 0, \quad R_\eta(t) = \sum_{n=0}^{\infty} a_{n+t}\overline{a}_n. \tag{3}$$

Quelles suites peut-on obtenir de la sorte?

LEMME 1. *Pour qu'une suite stationnaire $\eta(n)$ soit réponse d'un filtre physiquement réalisable à une suite standard de variables aléatoires, il faut et il suffit qu'elle possède une mesure spectrale absolu-*

ment continue et que sa densité spectrale $f(u)$ admette la représentation

$$f(u)=|g(e^{iu})|^2,\quad g(e^{iu})=\sum_{n=0}^{\infty} b_n e^{iun},\quad \sum_{n=0}^{\infty}|b_n|^2<\infty. \tag{4}$$

D é m o n s t r a t i o n. *Condition nécessaire.* Supposons que la suite se représente par (1). Posons

$$g(e^{iu})=\frac{1}{\sqrt{2\pi}}\sum_{n=0}^{\infty}\overline{a}_n e^{iun}. \tag{5}$$

La formule de Parseval nous apprend que

$$R_\eta(t)=\sum_{n=0}^{\infty} a_{n+t}\overline{a}_n=\int_{-\pi}^{\pi} e^{iut}\,|g(e^{iu})|^2\,du,$$

i.e. la suite $\eta(n)$ possède un spectre absolument continu de densité $f(u)=|g(e^{iu})|^2$.

Condition suffisante. Soit $\eta(n)$ une suite de fonction de corrélation

$$R_\eta(t)=\int_{-\pi}^{\pi} e^{iut}f(u)\,du$$

et $f(u)=|g(e^{iu})|^2$, où $g(e^{iu})$ est défini par (4). La suite $\eta(n)$ admet la représentation spectrale

$$\eta(n)=\int_{-\pi}^{\pi} e^{iun}\zeta(du).$$

Construisons la mesure stochastique

$$\xi(A)=\int_{-\pi}^{\pi}\frac{1}{\sqrt{2\pi}\,\overline{g(e^{iu})}}\,\chi_A(u)\,\zeta(du)$$

sur la tribu des boréliens de l'intervalle $[-\pi, \pi[$. On a

$$\mathsf{E}\xi(A)\overline{\xi(B)}=\int_{-\pi}^{\pi}\chi_A(u)\,\chi_B(u)\frac{1}{2\pi\,|g(e^{iu})|^2}f(u)\,du=\frac{1}{2\pi}\int_{A\cap B}du,$$

i.e. $\xi(A)$ est une mesure orthogonale, de fonction structurelle $l(A\cap B)$ où l est la mesure de Lebesgue. Les lemmes 2 et 1 montrent que

$$\eta(n)=\int_{-\pi}^{\pi} e^{iun}\zeta(du)=\int_{-\pi}^{\pi} e^{iun}\sqrt{2\pi}\,\overline{g(e^{iu})}\,\xi(du)=$$

$$= \sum_{k=0}^{\infty} \sqrt{2\pi}\, \overline{b}_k \int_{-\pi}^{\pi} e^{i(n-k)u} \xi(du) = \sum_{k=0}^{\infty} a_k \xi(n-k),$$

où

$$a_n = \sqrt{2\pi}\, \overline{b}_n, \quad \xi(n) = \int_{-\pi}^{\pi} e^{iun} \xi(du)$$

et

$$\mathsf{E}\xi(n)\overline{\xi(m)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-m)u}\, du = \delta_{nm}.$$

Donc $\xi(n)$ est une suite standard. ■

Le lemme démontré répond simplement à la question posée, mais dans le cas général cette réponse est insuffisante, car elle ne dit pas quand la densité spectrale admet la représentation (4).

Trouvons les conditions de représentabilité de $f(u)$ par (4). Soit H_2 l'ensemble des fonctions $f(z)$ analytiques dans le disque ouvert $D = \{z : |z| < 1\}$ et telles que

$$\| f(z) \|^2 = \lim_{r \uparrow 1} \int_{-\pi}^{\pi} |f(re^{i\theta})|^2\, d\theta < \infty.$$

Si $f(z) = \sum_{n=0}^{\infty} a_n z^n$, alors $f(re^{i\theta}) = \sum_{n=0}^{\infty} a_n r^n e^{in\theta}$, i.e. $a_n r^n$ sont les coefficients de Fourier de la fonction $f(re^{i\theta})$. L'égalité de Parseval nous montre que

$$\int_{-\pi}^{\pi} |f(re^{i\theta})|^2\, d\theta = 2\pi \sum_{n=0}^{\infty} |a_n|^2 r^{2n}.$$

D'où il vient que $f(z) \in H_2$ si et seulement si

$$\sum_{n=0}^{\infty} |a_n|^2 < \infty.$$

Donc pour toute fonction $f(z) \in H_2$ on peut définir une série $f(e^{i\theta}) = \sum_{n=0}^{\infty} a_n e^{in\theta}$ convergente dans $\mathcal{L}_2(l)$, où l est la mesure de Lebesgue sur $[-\pi, \pi[$. La fonction $f(z)$ $(|z| < 1)$ s'obtient à partir de la fonction $f(e^{i\theta})$ au moyen de la formule de

Poisson

$$f(re^{i\theta})=\frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{iu})\,P(r,\,\theta,\,u)\,du, \tag{6}$$

où

$$P(r,\,\theta,\,u)=\frac{1-r^2}{1-2r\cos(\theta-u)+r^2}=\sum_{n=-\infty}^{\infty} r^{|n|}e^{in(\theta-u)}.$$

Cette proposition est une conséquence immédiate de l'égalité de Parseval.

En théorie des fonctions (cf. I. Privalov [1]) on démontre que si dans la formule (6) la fonction $f(e^{i\theta})$ est intégrable-Lebesgue, alors pour presque tous les θ existe la

$$\lim_{r\uparrow 1} f(re^{i\theta}) = f(e^{i\theta}).$$

La fonction $f(e^{i\theta})$ est appelée *valeur frontière* de la fonction $f(z)$ $(|z|<1)$.

THÉORÈME 1. *Soit $f(u)$ une fonction non négative intégrable-Lebesgue sur $[-\pi,\pi[$. Pour qu'une fonction $g(z)\in H_2$ telle que*

$$f(u) = |g(e^{iu})|^2 \tag{7}$$

existe il est nécessaire et suffisant que

$$\int_{-\pi}^{\pi} |\ln f(u)|\,du<\infty. \tag{8}$$

Démonstration. *Condition nécessaire.* Soit $g(z)=\sum_{n=0}^{\infty} a_n z^n\in H_2$ et (7) a lieu. On peut admettre que $g(0)\neq 0$ (dans le cas contraire au lieu de $g(z)$ on considère $z^{-m}g(z)$, où m est la multiplicité du zéro ($z=0$) de la fonction $g(z)$, et on pose $|g(0)|=1$). Supposons que $0<r<1$ et $A=\{u:\ |g(re^{iu})|\leqslant 1\}$, $B=\{u:\ |g(re^{iu})|>1\}$. Alors

$$\int_{-\pi}^{\pi} |\ln|g(re^{iu})||\,du=\int_B \ln|g(re^{iu})|\,du-\int_A \ln|g(re^{iu})|\,du=$$

$$=2\int_B \ln|g(re^{iu})|\,du-\int_{-\pi}^{\pi} \ln|g(re^{iu})|\,du.$$

De la formule de Jensen il suit que pour $f(0) = 1$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln |f(re^{iu})|\, du = \ln \prod_{k=1}^{n} \frac{r}{|z_k|} \geqslant 0,$$

où z_k sont les zéros de la fonction $f(z)$ à l'intérieur du disque $|z| < r$. Donc

$$\int_{-\pi}^{\pi} |\ln |g(re^{iu})||\, du \leqslant 2\int_{B} \ln |g(re^{iu})|\, du \leqslant \int_{B} |g(re^{iu})|^2\, du \leqslant$$
$$\leqslant \int_{-\pi}^{\pi} |g(re^{iu})|^2\, du \leqslant 2\pi \sum_{n=0}^{\infty} |a_n|^2.$$

Le lemme de Fatou donne

$$\int_{-\pi}^{\pi} |\ln |g(e^{iu})||\, du = \int_{-\pi}^{\pi} \varliminf_{r\uparrow 1} |\ln |g(re^{iu})||\, du \leqslant$$
$$\leqslant \varliminf_{r\uparrow 1} \int_{-\pi}^{\pi} |\ln |g(re^{iu})||\, du \leqslant 2\pi \sum_{n=0}^{\infty} |a_n|^2,$$

ce qui démontre la nécessité de la condition (8).

Condition suffisante. Supposons la condition (8) réalisée. La fonction

$$v(r,\ \theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln f(u)\, P(r,\ \theta,\ u)\, du$$

est harmonique dans $D = \{z:\ |z| < 1\}$. L'inégalité de Jensen entraîne

$$v(r,\ \theta) \leqslant \ln\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} f(u)\, P(r,\ \theta,\ u)\, du\right\}.$$

Soit $\varphi(z)$ une fonction analytique dans D, de partie réelle $v(r,\ \theta)$. Posons $g(z) = e^{1/2\varphi(z)}$. Il vient

$$|g(re^{i\theta})|^2 = e^{\mathrm{Re}\,\varphi(z)} = e^{v(r,\ \theta)} \leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi} f(u)\, P(r,\ \theta,\ u)\, du$$

et

$$\int_{-\pi}^{\pi} |g(re^{i\theta})|^2\, d\theta \leqslant \int_{-\pi}^{\pi} |f(u)|\, du < \infty.$$

Donc

$$g(z) \in H_2$$

et

$$\lim_{r\uparrow 1} |g(re^{i\theta})|^2 = e^{\lim\limits_{r\uparrow 1} v(r,\, \theta)} = f(\theta)$$

presque partout.

Remarque 1. D'après la démonstration la fonction $g(z)$ peut être prise positive et sans zéros dans D pour $z = 0$.

Remarque 2. La fonction $g(z)$ dont l'existence est établie par le théorème 1 n'est pas définie de façon unique. Mais si $g(z)$ est telle que

a) $g(z) \neq 0,\ z \in D$;

b) $g(0) > 0$,

alors elle est unique et par suite est confondue avec celle que nous avons trouvée.

En effet, si $g_i(z)$ $(i = 1, 2)$ sont deux telles fonctions, $\psi(z) = \frac{g_1(z)}{g_2(z)}$ est analytique dans D, non nulle et égale à l'unité en module sur la frontière de D. La fonction $\ln \psi(z)$ est analytique dans D et de partie réelle nulle sur la frontière de D. Donc $\ln \psi(z) = ik$, où k est réel. Comme $\ln \psi(0)$ est réel, on déduit que $\ln \psi(z) = 0$.

Le lemme 1 et le théorème 1 entraînent le

Théorème 2. *Pour qu'une suite $\eta(t)$ soit représentable par*

$$\eta(t) = \sum_{n=0}^{\infty} a_n \xi(t-n), \qquad \sum_{n=0}^{\infty} |a_n|^2 < \infty,$$

où $\xi(n)$ est une suite non corrélée, il est nécessaire et suffisant que $\eta(t)$ possède une mesure spectrale absolument continue et une densité spectrale $f(u)$ telle que

$$\int_{-\pi}^{\pi} \ln f(u)\, du > -\infty.$$

Soient $\zeta_1(x)$, $\zeta_2(x)$, $x \in X$ deux fonctions aléatoires hilbertiennes. Appelons $\mathscr{L}_2\{\zeta_i\}$ l'enveloppe linéaire fermée du système de variables aléatoires $\{\zeta_i(x),\ x \in X\}$ dans $\mathscr{L}_2$.

Définition.. *Si $\mathscr{L}_2\{\zeta_1\} \subset \mathscr{L}_2\{\zeta_2\}$, la fonction aléatoire $\zeta_1(x)$ est dite adaptée à $\zeta_2(x)$. Si $\mathscr{L}_2\{\zeta_1\} = \mathscr{L}_2\{\zeta_2\}$, les fonctions $\zeta_1(x)$ et $\zeta_2(x)$ sont dites équivalentes.*

Remarque. L'équivalence des suites $\xi(\eta)$ et $\eta(n)$ découle de la démonstration du lemme 1.

Montrons comment dans la sommation glissante on peut exprimer les coefficients a_n en fonction de la densité spectrale $f(u)$ de la suite $\eta(t)$.

La fonction $\varphi(z)$ envisagée dans la démonstration du théorème 1 est une fonction analytique dans D, dont la partie réelle prend les valeurs frontières $\ln f(u)$. Donc, en vertu de la formule de Schwarz

$$\varphi(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln f(u) \frac{e^{iu}+z}{e^{iu}-z} du. \tag{9}$$

En développant la fonction $g(z) = \exp\left\{\frac{1}{2}\varphi(z)\right\}$ en série entière, $g(z) = \sum_{n=0}^{\infty} b_n z^n$, on obtient les valeurs suivantes pour les coefficients a_n :

$$a_n = \sqrt{2\pi}\, \overline{b}_n.$$

D'autre part, on peut transformer l'expression de $g(z)$ de la façon suivante. Comme

$$\frac{e^{iu}+z}{e^{iu}-z} = 1 + \frac{2ze^{-iu}}{1-ze^{-iu}} = 1 + 2\sum_{k=1}^{\infty} z^k e^{-iku},$$

il vient

$$\overline{g(z)} = \exp\left\{\frac{1}{4\pi} \int_{-\pi}^{\pi} \ln f(u)\, du + \frac{1}{2\pi} \sum_{k=1}^{\infty} d_k \overline{z}^k\right\},$$

où

$$d_k = \int_{-\pi}^{\pi} e^{iku} \ln f(u)\, du.$$

En posant

$$P = \exp\left\{\frac{1}{4\pi} \int_{-\pi}^{\pi} \ln f(u)\, du\right\}, \quad \exp\left\{\frac{1}{2\pi} \sum_{k=1}^{\infty} d_k z^k\right\} = \sum_{k=0}^{\infty} c_k z^k \quad (c_0 = 1),$$

on obtient

$$\overline{g(z)} = P \sum_{k=0}^{\infty} c_k \overline{z}^k,$$

et par suite

$$a_n \quad \sqrt{2\pi}\, P c_n. \tag{10}$$

Passons aux processus à temps continu. La généralisation de la sommation glissante aux processus aléatoires à temps continu est une opération associant à un processus aléatoire $\xi(t)$ le processus $\eta(t)$, $t \in]-\infty, \infty[$ défini par

$$\eta(t) = \int_0^\infty a(s)\, d\xi(t-s). \qquad (11)$$

Un processus $\xi(t)$ à accroissements orthogonaux est par définition standard si

$$\mathsf{E}\, \xi(t) = 0, \quad \mathsf{E}\, |\xi(t+h) - \xi(t)|^2 = h.$$

D'après ce qui a été dit au § 4 sur l'intégrale stochastique de Stieltjes, le processus $\xi(t)$ est associé à une mesure orthogonale stochastique $\xi(A)$ sur une tribu d'ensembles mesurables-Lebesgue. Cette mesure sera également appelée *mesure stochastique standard.* Pour que l'intégrale (11) existe il est nécessaire et suffisant que $a(t)$ soit mesurable-Lebesgue et

$$\int_0^\infty |a(t)|^2\, dt < \infty.$$

A noter que le processus standard $\xi(t)$ n'est pas dérivable en m.q. Cependant les quotients

$$\xi_\Delta^{\bullet}(t_k) = \frac{\xi(t_{k+1} + \Delta) - \xi(t_k)}{\Delta}, \quad \Delta = t_{k+1} - t_k,$$

sont orthogonaux pour tous les t_k et tous les Δ aussi petits que l'on veut. Il conviendrait donc de traiter la dérivée fictive $\xi'(t)$ comme un processus dont les valeurs en deux instants quelconques sont orthogonales et de variance infinie. Ce processus fictif intervient souvent dans les raisonnements sous le nom de *bruit blanc.* Le bruit blanc est défini exactement dans le cadre de la théorie des processus aléatoires généralisés (I. Guelfand et N. Vilenkine [1]). Symboliquement on peut mettre la formule (11) sous la forme

$$\eta(t) = \int_{-\infty}^{t} a(t-s)\, \xi'(s)\, ds$$

et interpréter $\eta(t)$ comme la réponse d'un filtre physiquement réalisable à un bruit blanc. La fonction de transfert impulsionnelle d'un tel filtre est nulle pour $t < 0$ et égale à $a(t)$ pour $t > 0$. A signaler que tous les filtres physiquement réalisables admissibles pour le processus $\xi'(t)$ sont donnés par la formule (11). En effet, tout filtre physiquement réalisable admissible, par définition, soit est de la forme (11), soit est limite d'un filtre de cette forme. La condition de convergence en m.q. de filtres de la forme (11), de fonctions de

transfert impulsionnelles $a_n(t)$, s'écrit:

$$\int_0^\infty |a_n(s) - a_{n'}(s)|^2 \, ds \to 0 \text{ pour } n, \; n' \to \infty.$$

Or, si cette condition est réalisée, l.i.m. $a_n(t) = a(t)$ existe (par rapport à la mesure de Lebesgue sur $]0, \infty[$) et

$$\text{l.i.m. } \eta_n(t) = \text{l.i.m.} \int_0^\infty a_n(s) \, d\xi(t-s) = \int_0^\infty a(s) \, d\xi(t-s).$$

Par suite le passage à la limite dans les filtres tels que (11) n'élargit pas la classe des filtres.

La formule (11) peut encore s'écrire:

$$\eta(t) = \int_{-\infty}^{\infty} a(t-s) \, d\xi(s), \quad a(t) = 0 \text{ pour } t < 0.$$

Donc la fonction de corrélation du processus $\eta(t)$ s'écrit

$$R(t) = \int_{-\infty}^{\infty} a(t+s-u) \overline{a(s-u)} \, du,$$

ou

$$R(t) = \int_0^\infty a(t+s) \overline{a(s)} \, ds. \tag{12}$$

LEMME 2. *Pour qu'un processus stationnaire (au sens large) $\eta(t)$ soit réponse d'un filtre physiquement réalisable à un bruit blanc attaché à $\eta(t)$ il est nécessaire et suffisant qu'il possède une mesure spectrale absolument continue et que sa densité spectrale $f(u)$ s'écrive*

$$f(u) = |h(iu)|^2, \tag{13}$$

où

$$h(iu) = \int_0^\infty b(s) e^{-ius} \, ds, \quad \int_0^\infty |b(s)|^2 \, ds < \infty. \tag{14}$$

Démonstration. *Condition nécessaire.* Supposons que le processus $\eta(t)$ admette la représentation (11). Posons

$$h(iu) = \frac{1}{\sqrt{2\pi}} \int_0^\infty a(s) e^{-ius} \, ds.$$

L'égalité de Parseval nous dit que

$$R(t) = \int_0^\infty a(t+s)\overline{a(s)}\,ds = \int_{-\infty}^\infty e^{iut}\,|h(iu)|^2\,du,$$

i.e. le spectre du processus est absolument continu et sa densité spectrale est de la forme (13), (14).

Condition suffisante. Soient réalisées les conditions du lemme. Considérons la représentation spectrale du processus $\eta(t)$:

$$\eta(t) = \int_{-\infty}^\infty e^{iut}\zeta(du),$$

et la mesure stochastique

$$\mu(A) = \int_{-\infty}^\infty \frac{\chi_A(u)}{h(iu)}\,\zeta(du). \tag{15}$$

L'intégrale stochastique (15) a un sens pour tout borélien borné A, puisque $\frac{\chi_A(u)}{h(iu)} \in \mathscr{L}_2\{F\}$, où F est mesure spectrale du processus

$$F(A) = \int_A |h(iu)|^2\,du.$$

Il est immédiat de remarquer que $\mu(A)$ est mesure orthogonale et que

$$\mathsf{E}\mu(A)\overline{\mu(B)} = \int_{A\cap B} du.$$

Posons

$$\xi(t_2) - \xi(t_1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{e^{-iut_2} - e^{-iut_1}}{-iu}\,\mu(du). \tag{16}$$

L'intégrale stochastique (16) existe visiblement. La fonction aléatoire d'intervalle $\xi(\Delta) = \xi(t_2) - \xi(t_1)$, $\Delta = [t_1, t_2[$ est une mesure élémentaire associée à un processus standard. En effet, $\mathsf{E}\xi(\Delta) = 0$. En appliquant ensuite l'égalité de Parseval aux intégrales de Fourier, on obtient

$$\mathsf{E}\xi(\Delta_1)\,\xi(\Delta_2) = \frac{1}{2\pi} \int_{-\infty}^\infty \frac{e^{-iut_2} - e^{-iut_1}}{-iu}\,\frac{e^{iut_4} - e^{iut_3}}{iu}\,du =$$

$$= \int_{-\infty}^\infty \chi_{\Delta_1}(t)\,\chi_{\Delta_2}(t)\,dt = l(\Delta_1 \cap \Delta_2),$$

où $\Delta_1 = [t_1, t_2[$, $\Delta_2 = [t_3, t_4[$, l est la mesure de Lebesgue sur la droite. Le lemme 1, § 2, et la formule (15) montrent que

$$\eta(t) = \int_{-\infty}^{\infty} e^{iut} h(iu)\, \mu(du). \tag{17}$$

On remarquera maintenant que si

$$h(iu) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} a(s)\, e^{ius}\, ds, \quad \text{où} \quad \int_{-\infty}^{\infty} |a(s)|^2\, ds < \infty,$$

alors

$$\int_{-\infty}^{\infty} h(iu)\, \mu(du) = \int_{-\infty}^{\infty} a(-s)\, \xi(ds). \tag{18}$$

En effet, les espaces $\mathscr{L}_2\{\mu\}$ et $\mathscr{L}_2\{\xi\}$ étant isomorphes à $\mathscr{L}_2\{l\}$, où l est la mesure de Lebesgue sur la droite $]-\infty, \infty[$, et la transformation de Fourier préservant le produit scalaire dans $\mathscr{L}_2\{l\}$, il suffit de vérifier la formule (18) pour des fonctions simples. Soit $a(t) = \sum c_k \chi_{\Delta_k}(t)$, où Δ_k est un intervalle ouvert (ou un intervalle semi-ouvert) $]a_k, b_k[$. Alors

$$\int_{-\infty}^{\infty} a(-s)\, \xi(ds) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sum_k c_k \frac{e^{iub_k} - e^{iua_k}}{iu}\, \mu(du)$$

est un cas particulier de (18). Nous avons donc établi la formule (18). De cette dernière il suit que

$$\int_{-\infty}^{\infty} e^{iut} h(iu)\, du = \int_{-\infty}^{\infty} a(s)\, d\xi(t-s), \tag{19}$$

puisque en vertu de (16) le produit de la mesure $\xi(\cdot)$ par e^{iut} équivaut à une t-translation de l'argument de la fonction $\xi(\cdot)$. De (17) et (19) il vient

$$\eta(t) = \int_{0}^{\infty} a(s)\, d\xi(t-s), \quad \text{où} \quad a(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(iu)\, e^{-iut}\, du. \quad \blacksquare$$

Soit donnée la densité spectrale $f(u)$ d'un processus $\eta(t)$. Les questions suivantes émergent : quand la densité spectrale admet-elle la représentation (ou factorisation) (13), (14) ? Comment calculer la fonction $h(iu)$ (et partant la fonction $a(t)$) sachant $f(u)$? On répond à ces questions en les ramenant à des questions déjà traitées par factorisation de fonctions sur le cercle. Soit la transformation $\omega = \frac{1+z}{1-z}$ associant au disque ouvert $D = \{z \colon |z| < 1\}$ le demi-

plan de droite $\Pi^+ = \{\omega : \text{Re}\,\omega > 0\}$. Sur la frontière de ces domaines ($\omega = iu$, $z=e^{i\theta}$) cette transformation s'écrit $u = \text{cotg}\,\frac{\theta}{2}$. Supposons que $f(u)$ admette la factorisation (13), (14). Posons

$$\left.\begin{aligned} g(z) &= (1+\omega)\,h(\omega) = \frac{2}{1-z}\,h\left(\frac{1+z}{1-z}\right), \\ \tilde{f}(\theta) &= f(u)\,(1+u^2). \end{aligned}\right\} \tag{20}$$

La fonction $\tilde{f}(\theta)$ admet la factorisation $|\tilde{f}(\theta)| = |g(e^{i\theta})|^2$, où $g(z)$ est analytique dans D et intégrable sur $]-\pi, \pi[$,

$$\int_{-\pi}^{\pi} \tilde{f}(\theta)\,d\theta = 2\int_{-\infty}^{\infty} f(u)\,du < \infty,$$

i.e. $g(z) \in H_2$. Le théorème 1 nous dit que

$$-\infty < \int_{-\pi}^{\pi} \ln \tilde{f}(\theta)\,d\theta = 2\int_{-\infty}^{\infty} \frac{\ln f(u) + \ln(1+u^2)}{1+u^2}\,du,$$

d'où

$$\int_{-\infty}^{\infty} \frac{\ln f(u)}{1+u^2}\,du > -\infty. \tag{21}$$

Supposons l'inverse. Soit $f(u)$ non négative, intégrable et vérifiant (21). Définissons $\tilde{f}(\theta)$ au moyen de (20). Alors $\tilde{f}(\theta)$ est intégrable et

$$\int_{-\pi}^{\pi} \ln \tilde{f}(\theta)\,d\theta > -\infty.$$

Le théorème 1 nous montre que $\tilde{f}(\theta)$ admet la factorisation

$$\tilde{f}(\theta) = |g(e^{i\theta})|^2,\ g(z) = \sum_0^{\infty} a_n z^n,\ \sum_{n=0}^{\infty} |a_n|^2 < \infty.$$

Posons

$$h(\omega) = \frac{1}{1+\omega}\sum_{n=0}^{\infty} a_n \left(\frac{1-\omega}{1+\omega}\right)^n.$$

La fonction $h(\omega)$ est alors analytique dans le demi-plan de droite et

$$f(u) = |h(iu)|^2,$$

où

$$h(iu) = \sum_{n=0}^{\infty} a_n \frac{(1-iu)^n}{(1+iu)^{n+1}}. \tag{22}$$

Vu que les fonctions $\frac{1}{\sqrt{2\pi}} e^{in\theta}$ forment une suite orthonormée complète dans $\mathscr{L}_2\ ]-\pi, \pi[$, il est immédiat de voir que la suite $\frac{1}{\sqrt{2\pi}} \frac{(1-iu)^n}{(1+iu)^{n+1}}$ l'est également sur $\mathscr{L}_2\ ]-\infty, \infty[$ par rapport à la mesure de Lebesgue. Donc la série (22) est convergente en m.q. On notera que

$$\frac{(1-iu)^n}{(1+iu)^{n+1}} = \sum_{k=1}^{n+1} \frac{A_k}{(1+iu)^k} =$$

$$= \sum_{k=1}^{n+1} \frac{A_k}{(k-1)!} \int_0^\infty e^{-(1+iu)\,t}\, t^{(k-1)}\, dt = \int_0^\infty e^{-iut} B_n(t)\, dt,$$

donc

$$h(iu) = \int_0^\infty e^{-iut} b(t)\, dt,$$

où

$$b(t) = \sum_{n=0}^{\infty} a_n B_n(t).$$

On rappelle que les sommes partielles de la série (22) sont les transformées de Fourier (à un facteur multiplicatif près) de fonctions égales à $\sum_1^N a_n B_n(t)$ pour $t \geqslant 0$ et nulles pour $t < 0$. La transformation de Fourier préservant la norme d'une fonction dans $\mathscr{L}_2\ ]-\infty, \infty[$, la convergence en m.q. de la série (22) implique celle de la série de $b(t)$ et

$$\int_0^\infty |b(t)|^2\, dt < \infty.$$

Au sujet de la factorisation de la fonction $f(u)$ on peut faire les mêmes remarques que pour la factorisation de fonctions sur un disque. L'expression de $h(\omega)$ se déduit de la formule (9) par substitution de ω à z et par un changement de variables approprié sous le signe de l'intégrale:

$$h(\omega) = \exp\left\{\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\ln f(u)}{1+u^2} \frac{i+u\omega}{u+i\omega}\, du\right\}. \tag{23}$$

THÉORÈME 3. *Pour qu'une fonction $f(u)$ $(-\infty < u < \infty)$ non négative intégrable admette la factorisation* (13), (14), *il est nécessaire*

et suffisant que

$$\int_{-\infty}^{\infty} \frac{\ln f(u)}{1+u^2}\, du > -\infty. \tag{24}$$

Si de plus l'on admet que $h(\omega) \neq 0$ $(\operatorname{Re} \omega > 0)$, $h(1) > 0$, *alors la fonction* $h(\omega)$ *est définie de façon unique par la formule* (23).

THÉORÈME 4. *Pour qu'un processus stationnaire* $\eta(t)$ $(-\infty < < t < \infty)$ *admette la représentation* (11) *il est nécessaire et suffisant que son spectre soit absolument continu et que sa densité spectrale vérifie la condition* (24).

§ 6. Prédiction et filtrage des processus stationnaires

En théorie des processus aléatoires il est un problème important, au champ d'applications vaste, qui consiste à trouver la meilleure estimation d'une variable aléatoire ζ sur le vu d'un ensemble de variables aléatoires $\{\xi_\alpha, \alpha \in A\}$. Il s'agit donc de trouver une fonction $f(\xi_\alpha \mid \alpha \in A)$ d'un ensemble de variables ξ_α, $\alpha \in A$, telle que

$$\zeta \approx \hat{\zeta} = f(\xi_\alpha \mid \alpha \in A) \tag{1}$$

avec la plus petite erreur.

Un problème de cette nature est la prédiction (extrapolation) d'un processus aléatoire. Il s'agit d'estimer ici la valeur d'un processus aléatoire à un instant t^* sachant les valeurs prises sur un ensemble d'instants antérieurs à t^*.

Un autre exemple est le filtrage d'un processus aléatoire. Dans le cas le plus simple ce problème se pose dans les termes suivants: un processus $\xi(t) = \eta(t) + \zeta(t)$, somme d'un signal « utile » $\zeta(t)$ et d'un « bruit » $\eta(t)$, est observé à des instants $t' \in T' \subset T$; on demande de séparer le bruit du signal, c'est-à-dire de trouver pour un $t^* \subset T$ la meilleure estimation

$$\zeta(t^*) \approx \hat{\zeta} = f(\xi(t') \mid t' \in T')$$

de $\zeta(t)$. Le problème sera entièrement posé lorsqu'on aura défini ce qu'on entend par « meilleure estimation ». Il va de soi que les critères d'optimalité dépendent de la nature du problème envisagé. La théorie mathématique développe essentiellement des méthodes de résolution basées sur l'écart quadratique moyen en tant que mesure de précision de l'égalité (1).

La quantité

$$\delta = \{E[\zeta - f(\xi_\alpha \mid \alpha \in A)]^2\}^{1/2} \tag{2}$$

s'appelle erreur quadratique moyenne de la formule approchée (1). Le problème est de définir une fonction f telle que (2) prenne une valeur minimale. Si A est un ensemble fini, par $f(\xi_\alpha \mid \alpha \in A)$ on entend une fonction mesurable-Borel de ξ_α, $\alpha \in A$. Si A par contre est infini, ce symbole désigne une variable aléatoire mesurable par rapport à la tribu $\mathfrak{F} = \sigma\{\xi_\alpha, \alpha \in A\}$ engendrée par l'ensemble de variables aléatoires $\{\xi_\alpha, \alpha \in A\}$.

On admet dans ce qui suit que ζ et $f(\xi_\alpha \mid \alpha \in A)$ possèdent des moments d'ordre deux.

Posons

$$\gamma = \mathsf{E}(\zeta \mid \mathfrak{F}). \tag{3}$$

Il vient

$$\delta^2 = \mathsf{E}\{\zeta - \hat{\zeta}\}^2 = \mathsf{E}(\zeta - \gamma)^2 + \\ + 2\mathsf{E}(\zeta - \gamma)(\gamma - \hat{\zeta}) + \mathsf{E}(\gamma - \hat{\zeta})^2.$$

La quantité $\gamma - \hat{\zeta}$ étant $\mathfrak{F}$-mesurable, on a

$$\mathsf{E}(\zeta - \gamma)(\gamma - \hat{\zeta}) = \mathsf{E}\mathsf{E}\{(\zeta - \gamma)(\gamma - \hat{\zeta}) \mid \mathfrak{F}\} = \\ = \mathsf{E}(\gamma - \hat{\zeta})\,\mathsf{E}\{(\zeta - \gamma) \mid \mathfrak{F}\} = 0.$$

Donc

$$\delta^2 = \mathsf{E}(\zeta - \gamma)^2 + \mathsf{E}(\gamma - \hat{\zeta})^2,$$

d'où suit le

Théorème 1. *L'estimation de ζ, variable aléatoire possédant un moment fini d'ordre deux, avec un écart quadratique moyen minimal, par $\hat{\zeta}$, variable aléatoire $\mathfrak{F} = \sigma\{\xi_\alpha, \alpha \in A\}$-mesurable, est unique* (mod P) *et est définie par*

$$\hat{\zeta} = \mathsf{E}\{\zeta \mid \mathfrak{F}\}.$$

Remarque. L'estimation $\hat{\zeta} = \gamma$ de la variable aléatoire ζ n'est pas biaisée, c'est-à-dire

$$\mathsf{E}\gamma = \mathsf{E}\mathsf{E}\{\zeta \mid \mathfrak{F}\} = \mathsf{E}\zeta,$$

et les quantités $\zeta - \gamma$ et ξ_α ne sont pas corrélées quel que soit $\alpha \in A$:

$$\mathsf{E}(\zeta - \gamma)\,\xi_\alpha = \mathsf{E}\mathsf{E}\{(\zeta - \gamma)\,\xi_\alpha \mid \mathfrak{F}\} = \mathsf{E}\xi_\alpha\ \mathsf{E}\{(\zeta - \gamma) \mid \mathfrak{F}\} = 0.$$

Malheureusement l'application du théorème 1 à l'obtention de formules efficaces d'estimation est une grande source de complication. Dans le cas de processus gaussiens on peut toutefois pousser nos raisonnements plus loin. Auparavant on remarquera qu'on obtiendrait souvent des solutions achevées et analytiquement accessibles en recherchant l'estimation optimale non pas sur la classe de toutes les fonctions mesurables des variables aléatoires données, mais sur une classe plus étroite de fonctions linéaires. Plus exactement, soit $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ l'espace probabilisé fondamental. Supposons que ξ_α et ζ possèdent des moments finis d'ordre deux. Soit $\mathscr{L}_2\{\xi_\alpha, \alpha \in A\}$

sous-espace de $\mathscr{L}_2\{\Omega, \mathfrak{S}, \mathsf{P}\}$, espace hilbertien et enveloppe linéaire fermée des variables ξ_α, $\alpha \in A$, et soient des constantes. Le sous-espace $\mathscr{L}_2\{\xi_\alpha, \alpha \in A\}$ peut être traité comme l'ensemble de toutes les fonctions linéaires (non homogènes) de ξ_α de variance finie. La meilleure estimation linéaire $\tilde{\zeta}$ de la variable aléatoire ζ est l'élément de $\mathscr{L}_2\{\xi_\alpha, \alpha \in A\}$ le plus proche de ζ, i.e.

$$\delta^2 = \mathsf{E}\,|\,\tilde{\zeta} - \zeta\,|^2 \leqslant \mathsf{E}\,|\,\zeta' - \zeta\,|^2$$

quel que soit $\zeta' \in \mathscr{L}_2\{\xi_\alpha, \alpha \in A\}$. En théorie des espaces hilbertiens il est classique que l'élément $\tilde{\zeta}$ du sous-espace H_0 le plus proche d'un élément donné ζ est unique. Plus exactement, $\tilde{\zeta}$ est la projection de ζ sur H_0. L'élément $\tilde{\zeta}$ est solution unique du système d'équations $(\zeta - \tilde{\zeta}, \zeta'') = 0$ pour tout $\zeta'' \in \mathscr{L}_2\{\xi_\alpha, \alpha \in A\}$. Dans notre cas ce système d'équations se ramène aux équations

$$\mathsf{E}\,(\tilde{\zeta}\,\bar{\xi}_\alpha) = \mathsf{E}\,(\zeta\bar{\xi}_\alpha), \tag{4}$$

et comme $\mathscr{L}_2\{\xi_\alpha, \alpha \in A\}$ contient l'unité,

$$\mathsf{E}\tilde{\zeta} = \mathsf{E}\zeta,$$

si bien que les estimations linéaires optimales $\tilde{\zeta}$ manifestement ne sont pas biaisées. On peut admettre que $\mathsf{E}\xi_\alpha = 0$ pour tout α. Dans la suite on se limitera donc à des sous-espaces de variables aléatoires de $\mathscr{L}_2\{\Omega, \mathfrak{S}, \mathsf{P}\}$ d'espérance mathématique nulle.

L'estimation linéaire de ζ n'est certes pas toujours valable. Si par exemple $\xi(n) = e^{i(\nu n+\varphi)}$, où ν est équi-réparti sur $]-\pi, \pi[$, alors $\mathsf{E}\{\xi(n)\,\overline{\xi(m)}\} = 0$ $(n \neq m)$ et la meilleure estimation linéaire de $\xi(m)$ sur les valeurs de $\xi(n)$ $(n \neq m)$ est de la forme $\tilde{\xi}(m) = 0$, c'est-à-dire n'utilise pas les valeurs de $\xi(n)$. D'autre part, un couple d'observations $\xi(k)$ et $\xi(k+1)$ suffit à déterminer exactement toute la suite $\xi(n)$, soit

$$\xi(n) = \left(\frac{\xi(k+1)}{\xi(k)}\right)^{n-k} \xi(k).$$

Supposons maintenant que toutes les répartitions finidimensionnelles du système $\{\zeta, \xi_\alpha, \alpha \in A\}$ sont normales et $\mathsf{E}\xi_\alpha = 0$, $\mathsf{E}\zeta = 0$. Dans ce cas les quantités $\zeta - \tilde{\zeta}$ et ξ_α sont indépen dantes, car non corrélées. Donc $\zeta - \tilde{\zeta}$ ne dépend pas de la tribu $\mathfrak{F}$ et

$$\mathsf{E}\{\zeta \mid \mathfrak{F}\} = \mathsf{E}\{\zeta - \tilde{\zeta} + \tilde{\zeta} \mid \mathfrak{F}\} = \mathsf{E}(\zeta - \tilde{\zeta}) + \tilde{\zeta} = \tilde{\zeta}.$$

THÉORÈME 2. *La meilleure estimation (au sens de l'écart quadratique moyen) d'une variable ζ d'un système gaussien $\{\zeta, \xi_\alpha, \alpha \in A\}$ par une fonction $\sigma\{\xi_\alpha, \alpha \in A\}$-mesurable est confondue avec la meilleure estimation linéaire sur $\mathscr{L}_2\{\xi_\alpha, \alpha \in A\}$.*

Dans la suite on traitera des problèmes particuliers de construction d'estimations linéaires optimales.

A) Les variables aléatoires ξ_α sont en nombre fini ($\alpha = 1, \ldots, n$). Le problème admet une solution simple, classique en algèbre linéaire. Si l'on admet que ξ_α sont linéairement indépendantes, on peut définir la projection $\tilde{\zeta}$ de ζ sur l'espace finidimensionnel H_0, tendu sur ξ_α ($\alpha = 1, \ldots, n$) par

$$\tilde{\zeta} = \frac{1}{G} \begin{vmatrix} (\xi_1, \xi_1) \ldots (\xi_n, \xi_1)\, \xi_1 \\ \ldots\ldots\ldots\ldots \\ (\xi_n, \xi_1) \ldots (\xi_n, \xi_n)\, \xi_n \\ (\zeta, \xi_1) \ldots (\zeta, \xi_n)\, 0 \end{vmatrix},$$

où $G = G(\xi_1, \xi_2, \ldots, \xi_n)$ est le déterminant de Gram des vecteurs $\xi_1, \xi_2, \ldots, \xi_n$:

$$G(\xi_1, \xi_2, \ldots \xi_n) = \begin{vmatrix} (\xi_1, \xi_1) \ldots (\xi_1, \xi_n) \\ \ldots\ldots\ldots\ldots \\ (\xi_n, \xi_1) \ldots (\xi_n, \xi_n) \end{vmatrix},$$

et $(\xi, \eta) = \mathsf{E}(\xi\bar{\eta})$. L'erreur quadratique moyenne δ de l'estimation $\zeta \approx \tilde{\zeta}$ est égale à la longueur de la perpendiculaire abaissée de l'extrémité du vecteur ξ sur l'espace H_0 et est définie par

$$\delta^2 = \frac{G(\xi_1, \xi_2, \ldots, \xi_n, \zeta)}{G(\xi_1, \xi_2, \ldots, \xi_n)}.$$

B) Estimation d'une variable aléatoire ζ sur le vu d'un processus aléatoire continu en m.q. $\xi(t)$ sur un intervalle de temps fini $T = [a, b]$. Soit $R(t, s)$ la fonction de corrélation du processus $\xi(t)$. Le théorème 5, § 1, nous dit que le processus $\xi(t)$ est développable en la série

$$\xi(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\varphi_k(t)\,\xi_k,$$

où $\varphi_k(t)$ est une suite orthonormée de fonctions propres, λ_k les valeurs propres de la fonction de corrélation sur $]a, b[$:

$$\lambda_k\varphi_k(t) = \int_a^b R(t, s)\,\varphi_k(s)\,ds,$$

ξ_k une suite normée non corrélée,

$$\mathsf{E}\xi_k\bar{\xi}_r = \delta_{kr}.$$

Les $\{\xi_k\}$, $k = 1, 2, \ldots$, forment visiblement une base dans $\mathcal{L}_2\{\xi(t),\ t \in\,]a, b[\}$. Donc

$$\tilde{\zeta} = \sum_{n=1}^{\infty} c_n\xi_n,$$

où

$$\sqrt{\lambda_n}\xi_n = \int_a^b \xi(t)\,\overline{\varphi_n(t)}\,dt,\ n = 1, 2, \ldots,$$

$$c_n = \mathsf{E}\zeta\bar{\xi}_n = \int_a^b R_{\zeta\xi}(t)\,\varphi(t)\,dt,\ R_{\zeta\xi}(t) = \mathsf{E}\zeta\bar{\xi}(t).$$

L'erreur quadratique moyenne δ est définie par

$$\delta^2 = \mathsf{E}\,|\,\zeta\,|^2 - \mathsf{E}\,|\,\tilde{\zeta}\,|^2 = \mathsf{E}\,|\,\zeta\,|^2 - \sum_{n=0}^{\infty}\left|\int_a^b R_{\zeta\xi}(t)\,\varphi_n(t)\,dt\right|^2.$$

L'usage de cette méthode est rendu malaisé par la lourdeur du calcul des fonctions et des valeurs propres du noyau $R\ (t,\ s)$.

C) **Méthode de Wiener.** Soient $\xi\ (t)$ et $\zeta\ (t)$, $t \in T$, deux fonctions aléatoires hilbertiennes. Supposons que le processus $\xi\ (t)$ est observé sur un ensemble T^* de valeurs de t. On demande la meilleure estimation de $\zeta\ (t_0)$, $t_0 \in T$, sur les valeurs observées de $\xi\ (t)$, $t \in T^*$. Si l'on admet que l'estimation cherchée est de la forme

$$\tilde{\zeta}(t_0) = \int_{T^*} c(s)\,\xi(s)\,m(ds), \tag{5}$$

où m est une mesure sur T^* et si sont réalisées les conditions sous lesquelles l'intégrale a un sens, alors l'équation (4) devient

$$\int_{T^*} c(s)\,R_{\xi\xi}(s,\ t)\,m(ds) = R_{\zeta\xi}(t_0,\ t),\quad t \in T^*, \tag{6}$$

où $R_{\xi\xi}$ est la fonction de corrélation de $\xi\ (t)$, et $R_{\zeta\xi}$ la fonction de corrélation de $\zeta\ (t)$ et $\xi\ (t)$. L'équation (6) est une équation intégrale de Fredholm de première espèce à noyau symétrique (hermitien). Elle n'admet pas toujours une solution. Cependant si

$$\int_T \mathsf{E}\,|\,\xi(t)\,|^2\,m(dt) < \infty,$$

elle admet une solution $c\ (s) \in \mathscr{L}_2\{m\}$ si et seulement si la meilleure estimation linéaire $\tilde{\zeta}\ (t_0)$ de $\zeta\ (t)$ est de la forme (5).

Si T est l'axe des réels, $T^* =]a,\ b[$, $\xi\ (t)$ et $\zeta\ (t)$ des processus stationnaires et stationnairement liés (au sens large) et m la mesure de Lebesgue, l'équation (6) s'écrit

$$\int_a^b c(s)\,R_{\xi\xi}(s - t)\,ds = R_{\zeta\xi}(t_0 - t),\quad t \in\,]a,\ b[. \tag{7}$$

Si $\zeta(t) = \xi(t)$ ($t \in]-\infty, \infty[$) et $t_0 > b$, i.e. le problème consiste à estimer $\xi(t_0)$ sur les valeurs de $\xi(t)$ dans le passé, ce problème est dit de *prédiction pure.*

Penchons-nous plus en détail sur la prédiction de $\zeta(t+q)$ sur le vu du processus $\xi(s)$ avant l'instant t, $t \geqslant s$. On supposera que les processus $\xi(t)$ et $\zeta(t)$ sont stationnaires et stationnairement liés (au sens large). La fonction $\tilde{\zeta}(t)$ aidant à la prédiction sera traitée comme une fonction de t à q fixe.

Posons

$$\tilde{\zeta}(t) = \int_{-\infty}^{t} c_t(s)\,\xi(s)\,ds.$$

Il est immédiat que le processus $\tilde{\zeta}(t)$ est stationnaire. En effet, l'équation (7) s'écrit

$$\int_{-\infty}^{t} c_t(s)\,R_{\xi\xi}(s-u)\,ds = R_{\zeta\xi}(t+q-u), \quad u \leqslant t.$$

Le changement de variables $t - u = v$, $t - s = \theta$ donne

$$\int_{0}^{\infty} c_t(t-\theta)\,R_{\xi\xi}(v-\theta)\,d\theta = R_{\zeta\xi}(q+v), \quad v \geqslant 0. \tag{8}$$

On voit que la fonction $c_t(t-\theta)$ ne dépend pas du temps, i.e. $c(\theta) = c_t(t-\theta)$. L'équation (8) devient

$$\int_{0}^{\infty} c(s)\,R_{\xi\xi}(t-s)\,ds = R_{\zeta\xi}(q+t), \quad t \geqslant 0, \tag{9}$$

et l'expression (5) de la fonction de prédiction s'écrit

$$\tilde{\zeta}(t) = \int_{-\infty}^{t} c(t-s)\,\xi(s)\,ds = \int_{0}^{\infty} c(s)\,\xi(t-s)\,ds. \tag{10}$$

Par suite le processus $\tilde{\zeta}(t) = \tilde{\zeta}_q(t)$ est stationnaire. La formule (10) dit que $c(t)$ est fonction de transfert impulsionnelle d'un filtre physiquement réalisable transformant le processus observé en la meilleure estimation de $\zeta(t+q)$.

Il est aisé de trouver l'expression de l'erreur quadratique moyenne δ de la fonction de prédiction $\tilde{\zeta}(t)$. Comme δ^2 est le carré de la perpendiculaire abaissée de l'extrémité du vecteur $\zeta(t+q)$ sur $\mathscr{L}_2\{\xi(s),\ s \leqslant t\}$, on a

$$\delta^2 = \mathsf{E}\,|\zeta(t+q)|^2 - \mathsf{E}\,|\tilde{\zeta}(t)|^2 =$$

$$= R_{\zeta\zeta}(0) - \int_{0}^{\infty}\int_{0}^{\infty} \overline{c(t)}\,R_{\xi\xi}(t-s)\,c(s)\,ds\,dt. \tag{11}$$

En admettant que $R_{\zeta\zeta}(0) = \sigma_\zeta^2$ et en passant à la représentation spectrale de la fonction de corrélation $R_{\xi\xi}(t)$, on obtient

$$\delta^2 = \sigma_\zeta^2 - \int_{-\infty}^{\infty} |c(iu)|^2 \, dF_{\xi\xi}(u), \tag{12}$$

où $F_{\xi\xi}(u)$ est la fonction spectrale du processus $\xi(t)$ et

$$c(iu) = \int_0^\infty c(t) e^{-iut} \, dt.$$

Exposons brièvement une méthode de résolution de l'équation (9) proposée par N. Wiener. Supposons que le spectre du processus $\xi(t)$ est absolument continu et que la densité spectrale $f_{\xi\xi}(u)$ admet la factorisation (cf. théorème 3, § 5)

$$f_{\xi\xi}(u) = |h(iu)|^2, \quad h(z) = \frac{1}{\sqrt{2\pi}} \int_0^\infty a(t) e^{-zt} \, dt, \quad \operatorname{Re} z \geqslant 0.$$

L'égalité de Parseval appliquée à la transformée de Fourier donne

$$R_{\xi\xi}(t) = \int_{-\infty}^{\infty} e^{iut} |h(iu)|^2 \, du = \int_0^\infty a(t+s) \overline{a(s)} \, ds.$$

Supposons encore que la fonction spectrale des processus $\zeta(t)$ et $\xi(t)$ est absolument continue et que sa densité $f_{\zeta\xi}(u)$ vérifie la condition

$$\frac{f_{\zeta\xi}(u)}{\overline{h(iu)}} = k(iu) \in \mathscr{L}_2. \tag{13}$$

On a alors

$$R_{\zeta\xi}(t) = \int_{-\infty}^{\infty} e^{iut} f_{\zeta\xi}(u) \, du = \int_{-\infty}^{\infty} e^{iut} k(iu) \overline{h(iu)} \, du = \int_0^\infty b(t+s) \overline{a(s)} \, ds,$$

où

$$b(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} k(iu) e^{iut} \, du.$$

L'équation (9) s'écrit maintenant

$$\int_0^\infty \left[b(q+t+s) - \int_0^\infty c(\theta) a(t-\theta+s) \, d\theta \right] \overline{a(s)} \, ds = 0, \quad t > 0. \tag{14}$$

Pour que (14) ait lieu il suffit que la fonction $c(t)$ soit solution de l'équation

$$b(q+x)=\int_0^\infty c(\theta)\,a(x-\theta)\,d\theta,\quad x>0. \tag{15}$$

L'équation (15) est du même type que l'équation (9) à la seule différence essentielle que la fonction $a(t)$ est nulle au-dessous de zéro. En mettant (15) sous la forme

$$b(q+x)=\int_0^x c(\theta)\,a(x-\theta)\,d\theta,\quad x>0, \tag{16}$$

on en obtient immédiatement la solution par la transformation de Laplace. En multipliant (16) par e^{-zx} et en intégrant entre 0 et ∞, on trouve

$$B_q(z)=C(z)\,h(z),$$

où

$$B_q(z)=\frac{1}{\sqrt{2\pi}}\int_0^\infty b(q+x)\,e^{-zx}\,dx,\quad C(z)=\frac{1}{\sqrt{2\pi}}\int_0^\infty c(t)\,e^{-zt}\,dt.$$

Par suite

$$C(z)=\frac{B_q(z)}{h(z)},\quad c(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{iut}\,\frac{B_q(iu)}{h(iu)}\,du, \tag{17}$$

l'expression de $B_q(z)$, $\operatorname{Re} z>0$, étant

$$B_q(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{iqu}\,\frac{f_{\zeta\xi}(u)}{\overline{h(iu)}}\,\frac{du}{z-iu}. \tag{18}$$

L'énoncé des hypothèses imposées aux formules (17), (18) est d'une grande lourdeur. Il est bien plus simple de vérifier immédiatement dans des problèmes concrets la validité des transformations qui donneraient éventuellement la solution.

D) **Méthode de Yaglom.** Dans la méthode de Yaglom, contrairement à celle de Wiener, on ne cherche pas la fonction de transfert impulsionnelle, qui est susceptible de ne pas exister, mais la caractéristique fréquentielle. A défaut de formules générales pour la résolution du problème on donne une méthode de détermination de la fonction cherchée par tâtonnement à partir des hypothèses qu'elle doit remplir. Dans nombre de cas importants ce choix est simple à réaliser.

Supposons qu'un processus stationnaire à deux dimensions $(\xi(t), \zeta(t))$ admet la représentation spectrale

$$\xi(t) = \int\limits_{-\infty}^{\infty} e^{iut} \nu_1(du), \quad \zeta(t) = \int\limits_{-\infty}^{\infty} e^{iut} \nu_2(du)$$

et la matrice de densité spectrale

$$\begin{pmatrix} f_{\xi\xi}(u) & f_{\xi\zeta}(u) \\ f_{\zeta\xi}(u) & f_{\zeta\zeta}(u) \end{pmatrix}.$$

On considère toujours le problème de la meilleure estimation de $\zeta(t+q)$ sur les valeurs du processus $\xi(s)$, $s \leqslant t$. Le processus de prédiction $\tilde{\zeta}(t)$ est attaché à $\xi(t)$. Donc

$$\tilde{\zeta}(t) = \int\limits_{-\infty}^{\infty} e^{iut} c(iu) \nu_1(du), \quad \int\limits_{-\infty}^{\infty} |c(iu)|^2 f_{\xi\xi}(u)\, du < \infty. \tag{19}$$

L'équation

$$\mathsf{E}\zeta(t+q)\overline{\xi(s)} = \mathsf{E}\tilde{\zeta}(t)\overline{\xi(s)}, \qquad s \leqslant t,$$

qui détermine le processus $\tilde{\zeta}(t)$ s'écrit

$$\int\limits_{-\infty}^{\infty} e^{ius} \{e^{iuq} f_{\zeta\xi}(u) - c(iu) f_{\xi\xi}(u)\}\, du = 0, \quad s > 0. \tag{20}$$

Outre les conditions (19) et (20), la fonction $c(iu)$ doit être caractéristique fréquentielle d'un filtre physiquement réalisable. Ces conditions seront réunies si

a) la fonction $f_{\xi\xi}(u)$ est bornée;

b) $c(iu)$ est valeur frontière de la fonction $c(z) \in \mathscr{H}_2^+$;

c) $\psi(iu) = e^{iuq} f_{\zeta\xi}(u) - c(iu) f_{\xi\xi}(u)$ est valeur frontière de la fonction $\psi(z)$ de $\mathscr{H}_2^-$.

$\mathscr{H}_2^+$ (resp. $\mathscr{H}_2^-$) désigne l'espace des fonctions $h(z)$ analytiques dans le demi-plan de droite (resp. de gauche) et telles que l'intégrale

$$\int\limits_{-\infty}^{\infty} |h(x+iu)|^2\, du$$

est uniformément bornée pour $x > 0$ (resp. $x < 0$).

En effet, b) implique $\int\limits_{\infty}^{\infty} |c(iu)|^2\, du < \infty$, ce qui avec a) entraîne la condition (19). D'autre part, de b) il suit que $c(iu)$ est caractéristique fréquentielle d'un filtre physiquement réalisable. La condition c) dit que $e^{iuq} f_{\zeta\xi}(u) - c(iu) f_{\xi\xi}(u)$ est la transformée de Fourier d'une fonction nulle au-dessus de zéro.

A remarquer qu'en ne retenant que l'hypothèse b) on élimine les filtres dont les caractéristiques croissent à l'infini. Ces caractéristiques fréquentielles sont associées à des opérations liées à la dérivation du processus $\xi(t)$ et se rencontrent souvent dans la construction de filtres optimaux. Il est préférable de remplacer la condition b) par une moins restrictive. Supposons que $c(z)$ est une fonction analytique dans le demi-plan de droite et $|c(z)| \to \infty$ avec $|z|$ pas plus vite qu'une puissance de z (par exemple, la puissance r). La fonction

$$c_n(z) = \frac{c(z)}{\left(1+\frac{z}{n}\right)^{r+1}} \in \mathscr{H}_2^+.$$

Comme $|c_n(z)| \leqslant |c(z)|$, il vient

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} |c_n(iu) - c(iu)|^2 f_{\xi\xi}(u)\,du = 0$$

sous réserve que (19) soit réalisée. Par suite, $c(iu)$ est limite dans $\mathscr{L}_2\{F_{\xi\xi}\}$ de caractéristiques fréquentielles de filtres physiquement réalisables, donc $c(iu)$ est également caractéristique fréquentielle d'un tel filtre. D'où le

THÉORÈME 3. *Si la densité spectrale* $f_{\xi\xi}(u)$ *d'un processus* $\xi(t)$ *est bornée, la caractéristique fréquentielle* $c(iu)$ *du filtre optimal estimant* $\zeta(t+q)$ *est définie de façon unique par les conditions*:

a) $$\int_{-\infty}^{\infty} |c(iu)|^2 f_{\xi\xi}(u)\,du < \infty,$$

b) $c(iu)$ *est valeur frontière de la fonction* $c(z)$, *analytique dans le demi-plan de droite et ne croît pas plus vite qu'une puissance de* $|z|$ *lorsque* $z \to \infty$,

c) $\psi(iu) = e^{iuq} f_{\zeta\xi}(u) - c(iu) f_{\xi\xi}(u)$ *est valeur frontière de la fonction* $\psi(z) \in \mathscr{H}_2^-$.

L'erreur quadratique moyenne δ sur la meilleure estimation vaut

$$\delta = \{\mathsf{E}\,|\zeta(t+q)|^2 - \mathsf{E}\,|\tilde{\zeta}(t)|^2\}^{1/2} =$$

$$= \left\{\sigma_\zeta^2 - \int_{-\infty}^{\infty} |c(iu)|^2 f_{\xi\xi}(u)\,du\right\}^{1/2}. \qquad (21)$$

EXEMPLE 1. Soit un processus de prédiction pure $\xi(t)$ ($\xi(t) = \zeta(t)$) de fonction de corrélation $R(t) = \sigma^2 e^{-\alpha|t|}$ ($\alpha > 0$). La densité spectrale se trouve sans peine: $f_{\xi\xi}(u) = \frac{\sigma^2\alpha}{\pi}\frac{1}{u^2+\alpha^2}$. Le prolongement analytique de la fonction $\psi(iu)$ a pour expression

$$\psi(z) = \frac{c(z) - e^{zq}}{(z+\alpha)(z-\alpha)}\frac{\sigma^2\alpha}{\pi}.$$

La fonction $\psi(z)$ admet un pôle unique $z = -\alpha$ dans le demi-plan de gauche. Pour l'éliminer au moyen d'une fonction $c(z)$ analytique dans le demi-plan de droite, il suffit de poser $c(z) = \text{const} = e^{-\alpha q}$. Ceci ne viole pas la condition a) du théorème 3. Ainsi

$$c(iu) = e^{-\alpha q}, \quad \tilde{\xi}(t) = \int_{-\infty}^{\infty} e^{iut} e^{-\alpha q} \nu(du),$$

i.e. la meilleure formule de prédiction de $\xi(t+q)$ est

$$\xi(t+q) \approx e^{-\alpha q} \xi(t),$$

formule qui ne dépend que de la valeur de $\xi(t)$ au dernier instant d'observation. L'erreur quadratique moyenne de l'extrapolation vaut

$$\delta = \sigma \sqrt{1 - e^{-2\alpha q}}.$$

EXEMPLE 2. On considère toujours un problème de prédiction pure d'un processus $\xi(t)$, i.e. l'estimation de $\xi(t+q)$ sur les valeurs observées $\xi(s)$, $s < t$. Si le spectre du processus $\xi(t)$ est absolument continu et est réalisée la condition (24), § 5, alors la densité spectrale admet la factorisation:

$$f_{\xi\xi}(u) = |h(iu)|^2,$$

où $h(z) \in \mathcal{H}_2^+$ ne possède pas de zéros dans le demi-plan de droite.

Etudions le cas important où $h(z)$ est une fraction rationnelle:

$$h(z) = \frac{P(z)}{Q(z)},$$

où $P(z)$ et $Q(z)$ sont des polynômes respectivement de degré m et n $(m < n)$. Supposons que la densité spectrale $f_{\xi\xi}(u)$ est bornée et ne s'annule pas. Les zéros des polynômes $P(z)$ et $Q(z)$ sont alors contenus dans le demi-plan de gauche. Soit

$$P(z) = A \prod_{j=1}^{p} (z - z_j)^{\alpha_j}, \quad Q(z) = B \prod_{j=1}^{r} (z - \tilde{z}_j)^{\beta_j},$$

$$\sum_{j=1}^{p} \alpha_j = m, \quad \sum_{j=1}^{r} \beta_j = n.$$

Posons

$$P_1(z) = (-1)^m \bar{A} \prod_{j=1}^{p} (z + \tilde{z}_j)^{\alpha_j}, \quad Q_1(z) = (-1)^n \bar{B} \prod_{j=1}^{r} (z + \tilde{\tilde{z}}_j)^{\beta_j}.$$

L'expression du prolongement analytique de la fonction $\psi(iu)$ est:

$$\psi(z) = (e^{zq} - c(z)) \frac{P(z)}{Q(z)} \frac{P_1(z)}{Q_1(z)}.$$

La fonction $c(z)$ doit être analytique sur le demi-plan de droite, $\psi(z)$ sur le demi-plan de gauche. Donc $c(z)$ doit l'être sur le plan tout entier de la variable complexe et est susceptible de présenter aux zéros du polynôme $P(z)$ des pôles de multiplicité non supérieure à celle du zéro correspondant. Par suite

$$c(z) = \frac{M(z)}{P(z)},$$

où $M(z)$ est une fonction analytique dans le plan z ne présentant pas de singularités pour z fini. $M(z)$ est un polynôme, car l'ordre de croissance de $c(z)$ n'est pas supérieur à l'ordre d'une puissance de z. L'intégrabilité du carré du module de la fonction

$$c(iu)\frac{P(iu)}{Q(iu)} = \frac{M(iu)}{Q(iu)}$$

entraîne que le degré m_1 du polynôme $M(iu)$ est $\leqslant n-1$.

Par ailleurs, la façon dont nous avons conduit le choix de la fonction $c(z)$ assure la réalisation des conditions a) et b) du théorème 3. Reste à prendre le polynôme $M(z)$ tel que la fonction

$$\psi(z) = \frac{[e^{zq}P(z) - M(z)]}{Q(z)}\frac{P_1(z)}{Q_1(z)}$$

ou, ce qui revient au même, que la fonction

$$\psi_1(z) = \frac{e^{zq}P(z) - M(z)}{Q(z)}$$

ne possède pas de pôles dans le demi-plan de gauche. Pour cela il faut et il suffit que

$$\left.\frac{d^jM(z)}{dz^j}\right|_{z=\tilde{z}_k} = \left.\frac{d^j(e^{zq}P(z))}{dz^j}\right|_{z=\tilde{z}_k}, \qquad (22)$$
$$j = 0, 1, \ldots, \beta_k - 1, \quad k = 1, \ldots, r.$$

La construction d'un polynôme $M(z)$ vérifiant les conditions (22) est un problème classique de la théorie de l'interpolation qui possède toujours une solution unique sur la classe des polynômes de degré $n-1$. La connaissance du polynôme $M(z)$ nous donne la caractéristique fréquentielle du filtre optimal de prédiction

$$c(iu) = \frac{M(iu)}{P(iu)}.$$

Voyons une autre méthode de détermination de la fonction $c(z)$. Décomposons les fonctions $P(z)\,Q^{-1}(z)$ et $M(z)\,Q^{-1}(z)$ en fractions élémentaires. Soit

$$\frac{P(z)}{Q(z)} = \sum_{k=1}^{r}\sum_{j=1}^{\beta_k}\frac{c_{kj}}{(z-\tilde{z}_k)^j}, \quad \frac{M(z)}{Q(z)} = \sum_{k=1}^{r}\sum_{j=1}^{\beta_k}\frac{\gamma_{kj}}{(z-\tilde{z}_k)^j}.$$

Pour que la fonction $\psi_1(z)$ ne possède pas de pôles aux points $\tilde{z}_k$, $k = 1, \ldots, r$, il est nécessaire et suffisant que

$$\frac{d^j}{dz^j}(z-\tilde{z}_k)^{\beta_k}\psi_1(z)\Big|_{z=\tilde{z}_k}=0,\ j=0,\ 1,\ \ldots,\beta_k-1,$$

et de plus

$$\psi_1(z)=\sum_{k=1}^{r}\sum_{j=1}^{\beta_k}\frac{c_{kj}e^{zq}-\gamma_{kj}}{(z-\tilde{z}_k)^j}.$$

Un calcul simple montre que

$$\gamma_{kj}=\left[c_{kj}+\frac{q}{1!}c_{kj+1}+\frac{q^2}{2!}c_{kj+2}+\ldots+\frac{q^{\beta_k-j}}{(\beta_k-j)!}c_{k\beta_k}\right]e^{\tilde{z}_k q}.$$

La connaissance des coefficients γ_{kj} nous permet d'écrire l'expression de $c(iu)$:

$$c(iu)=\frac{1}{h(iu)}\sum_{k=1}^{r}\sum_{j=1}^{\beta_k}\frac{\gamma_{kj}}{(z-\tilde{z}_k)^j}=\frac{\displaystyle\sum_{k=1}^{r}\sum_{j=1}^{\beta_k}\frac{\gamma_{kj}}{(z-\tilde{z}_k)^j}}{\displaystyle\sum_{k=1}^{r}\sum_{j=1}^{\beta_k}\frac{c_{kj}}{(z-\tilde{z}_k)^j}}.$$

Exemple 3. Supposons qu'on observe un processus $\zeta(s)$ $(s \leqslant t)$, mais que les résultats des mesures de $\zeta(s)$ sont faussés par des perturbations diverses de sorte que les valeurs observées donnent une fonction $\xi(s)$, $s \leqslant t$, différente de $\zeta(s)$. On admettra que la valeur des perturbations (ou encore du bruit) $\eta(t) = \xi(t) - \zeta(t)$ est un processus stationnaire d'espérance mathématique nulle. On demande l'estimation de $\zeta(t + q)$ sur les résultats des observations du processus $\xi(s) = \zeta(s) + \eta(s)$, $s \leqslant t$.

Ces problèmes sont dits problèmes de filtrage ou de lissage (on dit que le processus $\xi(t)$ doit être séparé du bruit $\eta(t)$ ou encore qu'il doit être « lissé », i.e. il faut en soustraire le bruit irrégulier). Si $q > 0$, on a affaire à un filtrage avec prédiction, et si $q < 0$, à un filtrage à retardement.

Supposons que le bruit $\eta(t)$ et le processus $\zeta(t)$ ne sont pas corrélés et possèdent les densités spectrales $f_{\xi\xi}(u)$ et $f_{\eta\eta}(u)$. On a alors

$$R_{\xi\xi}(t) = R_{\eta\eta}(t) + R_{\zeta\zeta}(t), \quad f_{\xi\xi}(u) = f_{\eta\eta}(u) + f_{\zeta\zeta}(u).$$

Comme $R_{\zeta\xi}(t) = R_{\zeta\zeta}(t)$, les processus $\zeta(t)$ et $\xi(t)$ admettent une densité spectrale commune et $f_{\zeta\xi}(u) = f_{\zeta\zeta}(u)$. Soient

$$f_{\zeta\zeta}(u)=\frac{c_1}{u^2+\alpha^2}, \quad f_{\eta\eta}(t)=\frac{c_2}{u^2+\beta^2}.$$

Alors

$$f_{\xi\xi}(u) = \frac{c_3(u^2+\gamma^2)}{(u^2+\alpha^2)(u^2+\beta^2)}, \quad c_3 = c_1 + c_2, \quad \gamma^2 = \frac{c_2\alpha^2 + c_1\beta^2}{c_1+c_2}.$$

La fonction $\psi(z)$ a pour expression

$$\psi(z) = \frac{-c_1 e^{zq}(z^2-\beta^2) + c_3 c(z)(z^2-\gamma^2)}{(z^2-\alpha^2)(z^2-\beta^2)}.$$

Soit $q > 0$. La fonction $\psi(z)$ doit être analytique dans le demi-plan de gauche et appartenir à $\mathcal{H}_2^-$. Pour que ceci ait lieu il faut que le numérateur s'annule aux points $z = -\alpha$ et $z = -\beta$. Ce qui nous conduit aux égalités

$$c(-\beta) = 0, \quad c(-\alpha) = \frac{c_1}{c_3} \frac{e^{-\alpha q}(\alpha^2-\beta^2)}{\alpha^2-\gamma^2}. \tag{23}$$

D'autre part, $c(z)$ doit être analytique dans le demi-plan de gauche (et d'après l'hypothèse b) dans celui de droite) sauf au point $z = -\gamma$ où elle peut posséder un pôle simple. Donc

$$c(z) = \frac{\varphi(z)}{z+\gamma},$$

où $\varphi(z)$ est une fonction entière. La finitude de l'intégrale

$$\int_{-\infty}^{\infty} |c(iu)|^2 f_{\xi\xi}(u)\, du$$

entraîne que $\varphi(z)$ est une fonction linéaire, $\varphi(z) = Az + B$.

De (23) il suit

$$c(z) = A\frac{z+\beta}{z+\gamma}, \quad A = \frac{c_1}{c_3}\frac{\beta+\alpha}{\gamma+\alpha} e^{-\alpha q}.$$

Donc la formule du lissage optimal avec prédiction s'écrit

$$\tilde{\zeta}_q(t) = A \int_{-\infty}^{\infty} e^{iut} \frac{iu+\beta}{iu+\gamma} \nu_1(du).$$

Comme $(iu+\gamma)^{-1}$ est caractéristique fréquentielle d'un filtre physiquement réalisable de fonction de transfert impulsionnelle $e^{-\gamma t}$, on obtient

$$\tilde{\zeta}_q(t) = \frac{c_1}{c_3}\frac{\beta+\alpha}{\gamma+\alpha} e^{-\alpha q} \left\{ \xi(t) + (\beta-\gamma) \int_{-\infty}^{t} e^{-\gamma(t-s)} \xi(s)\, ds \right\}. \tag{24}$$

La formule (24) n'est pas vraie pour $q < 0$. Formellement cela tient au fait que la fonction $\psi(z)$ n'est pas bornée dans le demi-plan de gauche. On peut la définir à partir des considérations suivantes. Soit

$$\psi_1(z) = -c_1 e^{zq}(z^2-\beta^2) + c_3 c(z)(z^2-\gamma^2).$$

La fonction $c(z)$ doit être analytique dans le demi-plan de gauche sauf aux points $z = -\gamma$ et $\psi_1(-\alpha) = \psi_1(-\beta) = 0$. Comme

$$c(z) = \frac{\psi_1(z) + c_1 e^{zq}(z^2 - \beta^2)}{c_3(z^2 - \gamma^2)}$$

et $c(z)$ est analytique dans le demi-plan de droite, il vient que $\psi_1(z)$ est une fonction entière et

$$\psi_1(\gamma) = -c_1 e^{\gamma q}(\gamma^2 - \beta^2). \tag{25}$$

Posons

$$\psi_1(z) = A(z)(z + \alpha)(z + \beta).$$

La fonction $A(z)$ doit être entière. De la condition a) du théorème 3 il suit que $A(z) = \text{const} = A$. La valeur de A se déduit à partir de l'équation (25):

$$A = c_1 e^{\gamma q} \frac{-\gamma + \beta}{\alpha + \gamma}.$$

D'où

$$c(iu) = \frac{c_1}{c_3} \frac{(\alpha+\gamma)(u^2+\beta^2) e^{iuq} - e^{\gamma q}(-\gamma+\beta)(iu+\alpha)(iu+\beta)}{(\alpha+\gamma)(u^2+\gamma^2)}. \tag{26}$$

La prédiction et le filtrage de suites stationnaires sont justiciables de méthodes analogues à celles exposées pour les processus à temps continu. Voyons un exemple.

Exemple 4. Soit une suite stationnaire $\xi(t)$ vérifiant l'équation simple d'autorégression

$$a_0\xi(t) + a_1\xi(t-1) + \ldots + a_p\xi(t-p) = \eta(t), \tag{27}$$

où $\eta(t)$ est une suite non corrélée standard et $\xi(t)$ est attaché à $\eta(t)$. Soient

$$\eta(t) = \int_{-\pi}^{\pi} e^{iut}\, d\zeta(u)$$

la représentation spectrale de la suite $\eta(t)$, $\zeta(u)$ un processus à accroissements non corrélés et de fonction structurelle $\frac{1}{2\pi} l(A \cap B)$, où l est la mesure de Lebesgue. La représentation spectrale de la suite $\xi(t)$ doit être de la forme

$$\xi(t) = \int_{-\pi}^{\pi} e^{iut}\varphi(u)\, d\zeta(u), \quad \text{où} \int_{-\pi}^{\pi} |\varphi(u)|^2\, du < \infty. \tag{28}$$

En portant (28) dans (27) on obtient

$$\int_{-\pi}^{\pi} e^{iut}\overline{P(e^{iu})}\,\varphi(u)\, d\zeta(u) = \int_{-\pi}^{\pi} e^{iut}\, d\zeta(u),$$

où $P(z)=\sum_{k=0}^{n}\overline{a}_k z^k$. D'où il suit

$$\varphi(u)=\frac{1}{\overline{P(e^{iu})}} \quad (\text{mod } l).$$

Supposons que la fonction $P(z)$ ne possède pas de zéros dans le disque fermé $|z|\leqslant 1$. Dans ces conditions, si

$$\frac{1}{P(z)}=\sum_{k=0}^{\infty}\overline{b}_k z^k \quad \left(b_0=\frac{1}{a_0}\right),$$

alors

$$\xi(t)=\sum_{n=0}^{\infty}b_n\eta(t-n),$$

i.e. nous avons obtenu la représentation de la suite $\xi(t)$ sous forme de réponse d'un filtre physiquement réalisable à la suite non corrélée $\eta(t)$. Comme

$$\xi(t)=-\frac{1}{a_0}[a_1\xi(t-1)+\ldots+a_p\xi(t-p)+\eta(t)], \tag{29}$$

la prédiction optimale de $\xi(t)$ sur les $\xi(t-n)$ $(n=1, 2, \ldots)$ donnés s'écrit

$$\tilde{\zeta}(t)=\frac{1}{a_0}[a_1\xi(t-1)+a_2\xi(t-2)+\ldots+a_p\xi(t-p)].$$

L'erreur quadratique moyenne minimale de prédiction vaut

$$\delta(t)=\left\{\mathsf{E}\frac{|\eta(t)|^2}{|a_0|^2}\right\}^{1/2}=\frac{1}{|a_0|}.$$

L'emploi réitéré de la formule (29) nous permet d'obtenir la prédiction optimale à long terme.

CHAPITRE VI

PROCESSUS À ACCROISSEMENTS INDÉPENDANTS

Les processus à accroissements indépendants ont déjà fait l'objet du § 4, chapitre I.

Dans ce chapitre on se propose d'étudier les propriétés des réalisations ainsi que diverses fonctionnelles de tels processus. Commençons par le cas le plus simple : la marche aléatoire qui est un processus homogène à temps discret.

§ 1. Marche aléatoire sur la droite

Soient $\xi_1, \xi_2, \ldots$ une suite de variables aléatoires indépendantes équi-réparties. Appelons *marche aléatoire* la suite des sommes

$$\zeta_0 = 0, \ \zeta_n = \xi_1 + \ldots + \xi_n \ (n = 1, 2, \ldots);$$

ζ_n est l'*état* de la particule errante à l'instant n, ξ_n le *n-ième pas* de la marche. On étudie parfois une marche issue d'un point x; dans ce cas $\zeta_0 = x$, $\zeta_n = x + \xi_1 + \ldots + \xi_n$. Dans ce paragraphe on envisagera des marches issues exclusivement du point 0. On admettra d'autre part que les ξ_k sont entiers. La marche est dite alors à *valeurs entières*.

On appelle *pas du réseau* de la marche le nombre entier maximal h tel que $\frac{1}{h}\, \xi_1$ soit également entier. Si $h = 1$, la promenade est dite *non réticulée*. De toute évidence, h est confondu avec le P.G.C.D. des k tels que $\mathsf{P}\{\xi_1 = k\} > 0$.

Notons $\varphi(z)$ la fonction caractéristique des variables ξ_k.

Lemme 1. *Si $z_0 = 0$ est la plus petite racine positive de l'équation $1 - \varphi(z) = 0$, alors $hz_0 = 2\pi$, où h est le pas du réseau de la marche.*

Démonstration. Posons:

$$p_k = \mathsf{P}\{\xi_1 = k\}, \ k = 0, \pm 1, \pm 2, \ldots$$

On trouve

$$0 = 1 - \varphi(z_0) = \sum p_k (1 - e^{ikz_0}) = \sum p_k (1 - \cos kz_0) = \sum p_k 2 \sin^2 \frac{kz_0}{2}.$$

Soit $N'=\{k:\ p_k>0\}$. Le nombre $\frac{kz_0}{2\pi}$ est alors entier pour $k\in N'$. Si h est le P.G.C.D. des nombres de N', alors $\frac{hz_0}{2\pi}$ est aussi entier, car

$$h=m_1k_1+\ldots+m_nk_n,$$

où m_i sont des entiers, $k_i\in N'$. Donc

$$\frac{hz_0}{2\pi}\geqslant 1,\quad z_0\geqslant\frac{2\pi}{h}.$$

D'autre part,

$$\varphi\left(\frac{2\pi}{h}\right)=\sum_{k\in N'}p_ke^{i\frac{2\pi}{h}k}=\sum_{k\in N'}p_k=1,$$

puisque $\frac{k}{h}$ est un entier pour $k\in N'$. Donc $\frac{2\pi}{h}$ est racine positive de l'équation $1-\varphi(z)=0$ et $z_0\leqslant\frac{2\pi}{h}$. ■

Corollaire. *Si la marche n'est pas réticulée, il existe* $\alpha>0$ *tel que*

$$\operatorname{Re}(1-\varphi(z))\geqslant\alpha z^2,\quad z\in[-\pi,\pi].$$

Soit $p_l\neq 0$. Il vient

$$\operatorname{Re}(1-\varphi(z))\geqslant 2p_l\sin^2\frac{lz}{2}\geqslant 2p_l\left(\frac{2}{\pi}\right)^2\left(\frac{lz}{2}\right)^2$$

si $\left|\frac{lz}{2}\right|\leqslant\frac{\pi}{2}$. Pour $\frac{\pi}{l}\leqslant|z|\leqslant\pi$ la fonction $\frac{z^2}{\operatorname{Re}(1-\varphi(z))}$ est continue (le lemme 1 dit que le numérateur n'est pas nul) et par suite bornée.

Il est immédiat de voir qu'il suffit d'étudier une marche non réticulée, puisque telle sera la suite $\left\{\frac{1}{h}\zeta_n,\ n=0,1,\ldots\right\}$, où h est le pas du réseau de la marche. Dans la suite on supposera donc que la marche est non réticulée.

Soit ν_x (x entier) variable aléatoire et instant de première arrivée de la marche aléatoire au point x:

$$\nu_x=\inf[n>0:\ \zeta_n=x]$$

(si $\zeta_n\neq x\ \forall n$, on admet que $\nu_x=+\infty$). La marche est *récurrente* si $\mathsf{P}\{\nu_0<+\infty\}=1$. Etudions les conditions de récurrence de la marche aléatoire. Il nous faudra à cet effet appliquer la transformation de Laplace à la variable ν_x.

Lemme 2. *Pour* $|\lambda|<1$

$$\mathsf{E}\lambda^{\nu_x}=\sum_{k=1}^{\infty}\lambda^k\mathsf{P}\{\nu_x=k\}=\left(\sum_{n=1}^{\infty}\lambda^n\mathsf{P}\{\zeta_n=x\}\right)\left(\sum_{n=0}^{\infty}\lambda^n\mathsf{P}\{\zeta_n=0\}\right)^{-1}$$

($\lambda^{+\infty}$ *est supposé nul*).

D é m o n s t r a t i o n. On a

$$\mathsf{P}\{\zeta_n = x\} = \mathsf{P}\{\zeta_n = x,\ \nu_x \leqslant n\} =$$
$$= \sum_{k=1}^{n} \mathsf{P}\{\zeta_n = x,\ \nu_x = k\} = \sum_{k=1}^{n} \mathsf{P}\{\zeta_n - \zeta_k = 0,\ \nu_x = k\}.$$

Il est évident que $\zeta_n - \zeta_k$ ne dépend pas de l'événement $\{\nu_x = k\}$. Donc

$$\mathsf{P}\{\zeta_n = x\} = \sum_{k=1}^{n} \mathsf{P}\{\nu_x = k\}\,\mathsf{P}\{\zeta_n - \zeta_k = 0\} =$$
$$= \sum_{k=1}^{n} \mathsf{P}\{\nu_x = k\}\,\mathsf{P}\{\zeta_{n-k} = 0\}.$$

En multipliant cette égalité par λ^n et en sommant sur n, on trouve

$$\sum_{n=1}^{\infty} \lambda^n \mathsf{P}\{\zeta_n = x\} = \sum_{k=1}^{\infty} \lambda^k \mathsf{P}\{\nu_x = k\} \cdot \sum_{n=0}^{\infty} \lambda^n \mathsf{P}\{\zeta_n = 0\}. \quad \blacksquare$$

COROLLAIRE. *Pour qu'une marche soit récurrente il est nécessaire et suffisant que*

$$\sum_{n=0}^{\infty} \mathsf{P}\{\zeta_n = 0\} = +\infty.$$

En effet,

$$\mathsf{P}\{\nu_0 < +\infty\} = \lim_{\lambda \uparrow 1} \sum_{k=1}^{\infty} \lambda^k \mathsf{P}\{\nu_0 = k\} =$$
$$= 1 - \lim_{\lambda \uparrow 1} \frac{1}{\sum_{n=0}^{\infty} \lambda^n \mathsf{P}\{\zeta_n = 0\}} = 1 - \frac{1}{\sum_{n=0}^{\infty} \mathsf{P}\{\zeta_n = 0\}}.$$

Formulons une condition de récurrence de la marche aléatoire au moyen de la fonction caractéristique d'un pas.

THÉORÈME 1. *Pour qu'une marche aléatoire non réticulée soit récurrente il est nécessaire et suffisant que*

$$\int_{-\pi}^{\pi} \operatorname{Re} \frac{1}{1 - \varphi(z)}\, dz = +\infty. \tag{1}$$

D é m o n s t r a t i o n. *Condition suffisante.* On a

$$\sum_{n=0}^{\infty} \lambda^n \mathsf{P}\{\zeta_n = 0\} = \sum_{n=0}^{\infty} \lambda^n \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathsf{E} e^{iz\zeta_n}\, dz = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{n=0}^{\infty} \lambda^n \varphi^n(z)\, dz =$$
$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{1 - \lambda\varphi(z)}\, dz = \frac{1}{2\pi} \operatorname{Re} \int_{-\pi}^{\pi} \frac{1}{1 - \lambda\varphi(z)}\, dz = \frac{1}{2\pi} \int_{-\pi}^{\pi} \operatorname{Re} \frac{1}{1 - \lambda\varphi(z)} dz.$$

Le théorème de Fatou donne

$$\sum_{n=0}^{\infty} \mathsf{P}\{\zeta_n=0\}=\lim_{\lambda\uparrow 1}\sum_{n=0}^{\infty}\lambda^n \mathsf{P}\{\zeta_n=0\}=\lim_{\lambda\uparrow 1}\frac{1}{2\pi}\int_{-\pi}^{\pi}\operatorname{Re}\frac{1}{1-\lambda\varphi(z)}\,dz\geqslant$$

$$\geqslant\frac{1}{2\pi}\int_{-\pi}^{\pi}\lim_{\lambda\uparrow 1}\frac{1}{1-\lambda\varphi(z)}\,dz=\frac{1}{2\pi}\int_{-\pi}^{\pi}\operatorname{Re}\frac{1}{1-\varphi(z)}\,dz.$$

La récurrence de la marche aléatoire résulte de (1) et du lemme 2.

Condition nécessaire. On raisonnera par l'absurde. On supposera que la marche est récurrente, mais que

$$\int_{-\pi}^{\pi}\operatorname{Re}\frac{1}{1-\varphi(z)}\,dz<\infty. \tag{2}$$

Soit $x\neq 0$. Etudions pour $0<\lambda<1$ l'expression

$$\sum_{n=0}^{\infty}\lambda^n\mathsf{P}\{\zeta_n=0,\ \nu_x>n\}=$$

$$=\sum_{n=0}^{\infty}\lambda^n\mathsf{P}\{\zeta_n=0\}-\sum_{n=0}^{\infty}\lambda^n\mathsf{P}\{\zeta_n=0,\ \nu_x\leqslant n\}=$$

$$=\sum_{n=0}^{\infty}\lambda^n\mathsf{P}\{\zeta_n=0\}-\sum_{n=0}^{\infty}\lambda^n\sum_{k=1}^{n}\mathsf{P}\{\nu_x=k\}\,\mathsf{P}\{\zeta_{n-k}=-x\}=$$

$$=\sum_{n=0}^{\infty}\lambda^n\mathsf{P}\{\zeta_n=0\}-\sum_{k=1}^{\infty}\lambda^k\mathsf{P}\{\nu_x=k\}\sum_{n=1}^{\infty}\lambda^n\mathsf{P}\{\zeta_n=-x\}.$$

En vertu du lemme 2 cette expression s'écrit encore

$$\sum_{n=0}^{\infty}\lambda^n\mathsf{P}\{\zeta_n=0\}(1-\mathsf{E}\lambda^{\nu_x}\mathsf{E}\lambda^{\nu_{-x}})=$$

$$=\sum_{n=0}^{\infty}\lambda^n\mathsf{P}\{\zeta_n=0\}(2-\mathsf{E}\lambda^{\nu_x}-\mathsf{E}\lambda^{\nu_{-x}}-(1-\mathsf{E}\lambda^{\nu_x})(1-\mathsf{E}\lambda^{\nu_{-x}})).$$

Comme $\mathsf{E}\lambda^{\nu_x}<1$, $\mathsf{E}\lambda^{\nu_{-x}}<1$, il vient

$$\sum_{n=0}^{\infty}\lambda^n\mathsf{P}\{\zeta_n=0,\ \nu_x>n\}\leqslant\sum_{n=0}^{\infty}\lambda^n\mathsf{P}\{\zeta_n=0\}(2-\mathsf{E}\lambda^{\nu_x}-\mathsf{E}\lambda^{\nu_{-x}})=$$

$$=\sum_{n=0}^{\infty}\lambda^n[2\mathsf{P}\{\zeta_n=0\}-\mathsf{P}\{\zeta_n=x\}-\mathsf{P}\{\zeta_n=-x\}]$$

(toujours en vertu du lemme 2). Par suite

$$\sum_{n=0}^{\infty} \lambda^n \mathsf{P}\{\zeta_n = 0,\ \nu_x > n\} \leqslant \sum_{n=0}^{\infty} \lambda^n \frac{1}{2\pi} \int_{-\pi}^{\pi} (2 - e^{-izx} - e^{izx})\, \varphi^n(z)\, dz =$$

$$= \frac{1}{\pi} \int_{-\pi}^{\pi} (1 - \cos zx) \frac{1}{1 - \lambda\varphi(z)}\, dz. \qquad (3)$$

Le corollaire du lemme 1 dit que la fonction

$$\frac{1 - \cos zx}{1 - \varphi(z)}$$

est bornée. D'autre part, la fonction $\frac{1-\varphi(z)}{1-\lambda\varphi(z)}$ est bornée pour $0 \leqslant \lambda \leqslant 1$ et $z \in [-\pi, \pi]$, puisque $|\varphi(z)| \leqslant 1$; la fonction $\frac{1-u}{1-\lambda u}$ est bornée si $\lambda \in [0, 1]$ et u varie dans le disque unitaire du plan complexe, car elle est continue si on la pose égale à l'unité pour $\lambda = 1$, $u = 1$. Donc, dans (3), on peut passer à la limite pour $\lambda \uparrow 1$:

$$\sum_{n=0}^{\infty} \mathsf{P}\{\zeta_n = 0,\ \nu_x > n\} \leqslant \frac{1}{\pi} \int_{-\pi}^{\pi} (1 - \cos zx) \frac{dz}{1 - \varphi(z)} \leqslant$$

$$\leqslant \frac{1}{\pi} \int_{-\pi}^{\pi} (1 - \cos zx)\, \mathrm{Re}\, \frac{1}{1 - \varphi(z)}\, dz.$$

De (2) il suit

$$\lim_{x \to \infty} \int_{-\pi}^{\pi} \cos zx\, \mathrm{Re}\, \frac{1}{1 - \varphi(z)}\, dz = 0$$

d'après le théorème de Riemann-Lebesgue. Par ailleurs, $\nu_x \to \infty$ pour $x \to \infty$ implique

$$\lim_{x \to \infty} \mathsf{P}\{\zeta_n = 0,\ \nu_x > n\} = \mathsf{P}\{\zeta_n = 0\}.$$

Donc pour tout m

$$\lim_{x \to \infty} \sum_{n=0}^{m} \mathsf{P}\{\zeta_n = 0,\ \nu_x > n\} = \sum_{n=0}^{m} \mathsf{P}\{\zeta_n = 0\} \leqslant \frac{1}{\pi} \int_{-\pi}^{\pi} \mathrm{Re}\, \frac{1}{1 - \varphi(z)}\, dz$$

et par suite

$$\sum_{n=0}^{\infty} \mathsf{P}\{\zeta_n = 0\} \leqslant \frac{1}{\pi} \int_{-\pi}^{\pi} \mathrm{Re}\, \frac{1}{1 - \varphi(z)}\, dz < \infty$$

(sous l'hypothèse (2)), ce qui contredit la récurrence de la marche. ■

Soit $x \geqslant 0$. Appelons τ_x l'instant de première entrée dans l'ensemble (x, ∞):

$$\tau_x = \inf [n : \zeta_n > x]$$

(si $\zeta_n \leqslant x \ \forall\ x$, on admet que $\tau_x = +\infty$). Si τ_x est fini, on pose $\gamma_x = \zeta_{\tau_x} - x$. On cherchera la répartition conjointe de τ_x et γ_x. Soit $|\lambda| < 1$, $z \in [-\pi, \pi]$. Posons

$$u_x(\lambda, z) = \mathsf{E} e^{i\gamma_x z} \lambda^{\tau_x} \chi_{\{\tau_x < \infty\}} = \sum_{k=1,\ m=1}^{\infty} e^{ikz} \lambda^m \mathsf{P}\{\tau_x = m,\ \gamma_x = k\}.$$

Soit $p_l = \mathsf{P}\{\xi_1 = l\}$. Pour $m > 1$ on a

$$\mathsf{P}\{\tau_x = m,\ \gamma_x = k\} = \mathsf{P}\{\zeta_1 \leqslant x,\ \ldots,\ \zeta_{m-1} \leqslant x,\ \zeta_m = k + x\} =$$

$$= \sum_{l \leqslant x} \mathsf{P}\{\zeta_1 \leqslant x,\ \ldots,\ \zeta_{m-1} \leqslant x,\ \zeta_m = k + x,\ \xi_1 = l\} =$$

$$= \sum_{l \leqslant x} p_l \mathsf{P}\{\zeta_2 - \zeta_1 \leqslant x - l,\ \ldots,\ \zeta_{m-1} - \zeta_1 \leqslant x - l,\ \zeta_m - \zeta_1 = k + x - l\} =$$

$$= \sum_{l \leqslant x} p_l \mathsf{P}\{\tau_{x-l} = m - 1,\ \gamma_{x-l} = k\}.$$

De plus

$$\mathsf{P}\{\tau_x = 1,\ \gamma_x = k\} = \mathsf{P}\{\zeta_1 = k + x\} = p_{k+x}.$$

Donc

$$\sum_{m=1,\ k=1}^{\infty} \lambda^m e^{ikz} \mathsf{P}\{\tau_x = m,\ \gamma_x = k\} =$$

$$= \lambda \sum_{k=1}^{\infty} e^{ikz} p_{k+x} + \sum_{l \leqslant x} p_l \sum_{m=2,\ k=1}^{\infty} \lambda^m e^{ikz} \mathsf{P}\{\tau_{x-l} = m - 1,\ \gamma_{x-l} = k\},$$

ou

$$u_x(\lambda, z) = \lambda g_x(z) + \lambda \sum_{l \leqslant x} p_l u_{x-l}(\lambda, z), \tag{4}$$

où

$$g_x(z) = \sum_{k=1}^{\infty} e^{ikz} p_{k+x}.$$

Indiquons une méthode de résolution du système d'équations

$$u_x = g_x + \lambda \sum_{l \leqslant x} p_l u_{x-l}, \tag{5}$$

où u_x et g_x sont définis pour $x \geqslant 0$ et bornés. Posons $u_x = 0$,

$g_x = -\lambda \sum_{l \leqslant x} p_l u_{x-l}$ pour $x < 0$. Le système (5) s'écrit

$$u_x = g_x + \lambda \sum p_l u_{x-l}, \quad -\infty < x < \infty. \tag{6}$$

Nous aurons besoin du lemme suivant.

LEMME 3. *Supposons que les c_k sont déterminés à partir de l'égalité*

$$\sum_{-\infty}^{\infty} c_k e^{ikz} = \exp\left\{\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k=-\infty}^{0} \mathsf{P}\{\zeta_n = k\} e^{ikz}\right\}. \tag{7}$$

Alors $c_k = 0$ pour $k > 0$, $\sum |c_k| \leqslant \frac{1}{1-|\lambda|}$ et

$$c_k = \lambda \sum_l p_l c_{k-l}, \quad k < 0. \tag{8}$$

Démonstration. La série entière en e^{iz} du second membre de (7) étant absolument convergente, en substituant u, $|u| \leqslant 1$, à e^{iz} on obtiendrait le premier membre de (7) par développement de $\exp\{\cdot\}$ en série entière et par identification des coefficients en les mêmes harmoniques. On constate alors que les coefficients en e^{ikz} pour $k > 0$ sont nuls et que

$$\sum |c_k| \leqslant \exp\left\{\sum_{n=1}^{\infty} \frac{|\lambda|^n}{n} \sum_{k=-\infty}^{0} \mathsf{P}\{\zeta_n = k\}\right\} \leqslant$$

$$\leqslant \exp\left\{\sum_{n=1}^{\infty} \frac{|\lambda|^n}{n}\right\} = \frac{1}{1-|\lambda|}.$$

Ensuite

$$\sum_{k=-\infty}^{0} c_k e^{ikz} = \exp\left\{\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k=-\infty}^{\infty} \mathsf{P}\{\zeta_n = k\} e^{ikz}\right\} \times$$

$$\times \exp\left\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k=1}^{\infty} \mathsf{P}\{\zeta_n = k\} e^{ikz}\right\} =$$

$$= \exp\left\{\sum_{n=1}^{\infty} \frac{\lambda^n \varphi^n(z)}{n}\right\} \exp\left\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k=1}^{\infty} \mathsf{P}\{\zeta_n = k\} e^{ikz}\right\} =$$

$$= \frac{1}{1-\lambda\varphi(z)} \exp\left\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k=1}^{\infty} \mathsf{P}\{\zeta_n = k\} e^{ikz}\right\}.$$

Donc

$$\left(1-\lambda \sum p_l e^{ilz}\right) \sum c_k e^{ikz} = \exp\left\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k=1}^{\infty} \mathsf{P}\{\zeta_n = k\} e^{ikz}\right\}. \qquad (9)$$

On établit d'une façon analogue que le second membre se décompose en une série de Fourier seulement sur e^{ikz} pour les $k > 0$. Par suite, pour $k < 0$

$$c_k - \lambda \sum p_l c_{k-l} = 0$$

(c'est le coefficient du premier membre en e^{ikz}). ■

Remarque. De l'expression (9) il résulte que pour $k \geqslant 0$

$$c_k - \lambda \sum p_l c_{k-l} = b_k, \qquad (10)$$

où b_k sont tirés de l'expression

$$\sum_{k=1}^{\infty} b_k e^{ikz} = \exp\left\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k=1}^{\infty} \mathsf{P}\{\zeta_n = k\} e^{ikz}\right\}. \qquad (11)$$

Revenons à l'équation (6). On a

$$\sum c_k u_{x-k} = \sum c_k g_{x-k} + \lambda \sum p_l \sum_k c_k u_{x-l-k},$$

ou

$$\sum c_k g_{x-k} = \sum \left(c_k - \sum_l p_l c_{k-l}\right) u_{x-k}.$$

Le lemme 3 et la remarque subséquente entraînent que

$$\sum_{k=-\infty}^{0} c_k g_{x-k} = \sum_{k=0}^{\infty} b_k u_{x-k}. \qquad (12)$$

On remarquera que le second membre s'annule pour $x < 0$, donc le premier s'annulera également. Soit maintenant $x \geqslant 0$. Dans $\sum\limits_{k=-\infty}^{0} c_k g_{x-k}$ il ne figure que les g à indices négatifs qui sont connus. En multipliant (12) par ρ^x, où $|\rho| < 1$, et en sommant sur x, on trouve

$$\sum_{x=0}^{\infty} \rho^x \sum_{k=-\infty}^{0} c_k g_{x-k} = \sum_{k=0}^{\infty} \rho^k b_k \sum_{x=0}^{\infty} \rho^x u_x.$$

Donc

$$\sum_{x=0}^{\infty} \rho^x u_x = \left(\sum_{k=0}^{\infty} \rho^k b_k\right)^{-1} \sum_{x=0}^{\infty} \rho^x \sum_{k=-\infty}^{0} c_k g_{x-k}.$$

On remarquera maintenant que la fonction

$$\exp\left\{-\sum_{n=1}^{\infty}\frac{\lambda^n}{n}\sum_{k=1}^{\infty}\mathsf{P}\{\zeta_n=k\}\rho^k\right\}$$

est analytique pour $|\rho|<1$ et continue pour $|\rho|\leqslant 1$. Comme pour $\rho=e^{iz}$

$$\sum_{k=0}^{\infty}e^{ikz}b_k=\exp\left\{-\sum_{n=1}^{\infty}\frac{\lambda^n}{n}\sum_{k=1}^{\infty}\mathsf{P}\{\zeta_n=k\}e^{izk}\right\},$$

on a

$$\sum_{k=0}^{\infty}\rho^k b_k=\exp\left\{-\sum_{n=1}^{\infty}\frac{\lambda^n}{n}\sum_{k=1}^{\infty}\mathsf{P}\{\zeta_n=k\}\rho^k\right\}. \tag{13}$$

Par suite

$$\sum_{x=0}^{\infty}\rho^x u_x=\exp\left\{\sum_{n=1}^{\infty}\frac{\lambda^n}{n}\sum_{k=1}^{\infty}\mathsf{P}\{\zeta_n=k\}\rho^k\right\}\sum_{x=0}^{\infty}\rho^x\sum_{k=-\infty}^{\infty}c_k g_{x-k}. \tag{14}$$

Tirons u_0 de (14). Posons à cet effet $\rho=0: u_0=\sum_{k=-\infty}^{\infty}c_k g_{-k}$. Le calcul de $u_0(\lambda, z)$ passe par la substitution de $\lambda g_{-k}(z)$ à g_{-k}:

$$u_0(\lambda, z)=\lambda\sum_{k=-\infty}^{\infty}c_k\sum_{j=1}^{\infty}p_{j-k}e^{ijz}.$$

Soit l'opération $[\ ]_+$ qui à toute série trigonométrique $\sum_k \alpha_k e^{ikz}$ associe la série trigonométrique $\sum_{k>0}\alpha_k e^{ikz}$. On a

$$u_0(\lambda, z)=\lambda\left[\sum_{k=-\infty}^{\infty}c_k\sum_{j=-\infty}^{\infty}p_{j-k}e^{ijz}\right]_+=\lambda\left[\sum_k e^{ikz}c_k\varphi(z)\right]_+=$$

$$=-\left[\sum_k c_k e^{ikz}(1-\lambda\varphi(z))\right]_+,$$

puisque $\left[\sum_k c_k e^{ikz}\right]_+=0$. La relation (9) et l'égalité

$$\left[\exp\left\{-\sum_{n=1}^{\infty}\frac{\lambda^n}{n}\sum_{k\geqslant 1}\mathsf{P}\{\zeta_n=k\}e^{ikz}\right\}-1\right]_+=$$

$$=\exp\left\{-\sum_{n=1}^{\infty}\frac{\lambda^n}{n}\sum_{k\geqslant 1}\mathsf{P}\{\zeta_n=k\}e^{ikz}\right\}-1$$

donnent

$$u_0(\lambda, z) = 1 - \exp\left\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k\geqslant 1} \mathsf{P}\{\zeta_n = k\} e^{ikz}\right\}. \tag{15}$$

Trouvons la fonction génératrice de τ_0 en faisant $z = 0$ dans (15):

$$\mathsf{E}\lambda^{\tau_0} = 1 - \exp\left\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \mathsf{P}\{\zeta_n > 0\}\right\}. \tag{16}$$

THÉORÈME 2. *Pour que la marche soit majorée, i.e.*

$$\mathsf{P}\{\sup_n \zeta_n < +\infty\} = 1, \tag{17}$$

il est nécessaire et suffisant que

$$\sum_{n=1}^{\infty} \frac{1}{n} \mathsf{P}\{\zeta_n > 0\} < \infty. \tag{18}$$

Démonstration. De (16) il suit que

$$\mathsf{P}\{\tau_0 = +\infty\} = 1 - \lim_{\lambda \uparrow 1} \mathsf{E}\lambda^{\tau_0} =$$

$$= \lim_{\lambda \uparrow 1} \exp\left\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \mathsf{P}\{\zeta_n > 0\}\right\} = \exp\left\{-\sum_{n=1}^{\infty} \frac{1}{n} \mathsf{P}\{\zeta_n > 0\}\right\}$$

(on admet que $e^{-\infty} = 0$). Si donc la condition (17) est réalisée, on a

$$\mathsf{P}\{\sup_n \zeta_n < +\infty\} \geqslant \mathsf{P}\{\tau_0 = +\infty\} > 0.$$

Or, en vertu de la loi de tout ou rien (cf. chapitre II, § 4, théorème 7), $\mathsf{P}\{\sup_n \zeta_n < +\infty\}$ ne peut prendre que la valeur 0 ou 1. Par suite, (17) est réalisée.

Soit $\mathsf{P}\{\tau_0 = +\infty\} = 0$. Il est équivalent d'écrire

$$\sum_{n=1}^{\infty} \frac{1}{n} \mathsf{P}\{\zeta_n > 0\} = +\infty.$$

Si $\mathsf{P}\{\xi_1 \geqslant 0\} = 1$, alors

$$\mathsf{P}\{\sup_n \zeta_n = +\infty\} = 1, \tag{19}$$

car $\sum_{k=1}^{\infty} \xi_k = +\infty$ presque sûrement. Supposons que $\mathsf{P}\{\xi_1 < 0\} > 0$ et soit $m < 0$ tel que $p_m > 0$. Pour tous les n on a alors

$$\mathsf{P}\{\xi_1 = m, \ldots, \xi_n = m, \sup_{k\geqslant 0} (\zeta_{n+k} - \zeta_n) \leqslant -nm\} \leqslant$$

$$\leqslant \mathsf{P}\{\sup \zeta_l \leqslant 0\} = \mathsf{P}\{\tau_0 = +\infty\} = 0.$$

Or le premier membre vaut

$$(p_m)^n \, \mathsf{P}\{\sup_k \zeta_k \leqslant -nm\}.$$

Donc $\mathsf{P}\{\sup_k \zeta_k < n\} = 0 \ \forall n$. D'où (19). ■

Utilisons maintenant la relation (14) pour déterminer la répartition de

$$\zeta^+ = \sup_{n \geqslant 0} \zeta_n.$$

De toute évidence,

$$\mathsf{P}\{\zeta^+ = m\} = \mathsf{P}\{\tau_{m-1} < \infty, \ \tau_m = +\infty\} = \\ = \mathsf{P}\{\tau_{m-1} < \infty\} - \mathsf{P}\{\tau_m < \infty\}.$$

Donc

$$\mathsf{P}\{\zeta^+ = m\} = u_{m-1}(1, 0) - u_m(1, 0).$$

Soit $0 < \rho < 1$. Il vient

$$\sum_{m=0}^{\infty} \mathsf{P}\{\zeta^+ = m\} \rho^m = \sum_{m=1}^{\infty} \rho^m [u_{m-1}(1, 0) - u_m(1, 0)] + \mathsf{P}\{\zeta^+ = 0\} = \\ = 1 - \mathsf{P}\{\tau_0 < \infty\} - \sum_{m=1}^{\infty} \rho^m u_m(1, 0) + \rho \sum_{m=0}^{\infty} \rho^m u_m(1, 0) = \\ = 1 + (\rho - 1) \sum_{m=0}^{\infty} \rho^m u_m(1, 0). \qquad (20)$$

Pour obtenir $\sum_{m=0}^{\infty} \rho^m u_m(1, 0)$ il faut poser $z = 0$ dans le second membre et passer ensuite à la limite pour $\lambda \uparrow 1$. Par définition, les nombres g_x vaudront

$$g_x = \lambda \sum_{k=1}^{\infty} p_{k+x}.$$

Donc

$$(\rho - 1) \sum_{m=0}^{\infty} \rho^m u_m(\lambda, 0) = \exp\left\{\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k=1}^{\infty} \mathsf{P}(\zeta_n = k) \rho^k)\right\} \times \\ \times \sum_{x=0}^{\infty} (\rho^{x+1} - \rho^x) \sum_{k=-\infty}^{\infty} c_k g_{x-k} = \exp\left\{\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k=1}^{\infty} \mathsf{P}\{\zeta_n = k\} \rho^k\right\} \times \\ \times \left[-\sum_{k=-\infty}^{\infty} c_k g_{-k} + \sum_{x=1}^{\infty} \rho^x \sum_{k=-\infty}^{\infty} c_k (g_{x-k-1} - g_{x-k})\right] =$$

$$= \exp\left\{\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k=1}^{\infty} \mathsf{P}\{\zeta_n = k\} \rho^k\right\} \times$$

$$\times \left[-u_0(\lambda, 0) + \lambda \sum_{x=1}^{\infty} \rho^x \sum_{k=-\infty}^{\infty} c_k p_{x-k} \right].$$

La remarque sur le lemme 3 et la formule (11) entraînent

$$\lambda \sum_{x=1}^{\infty} \rho^x \sum_{k=-\infty}^{\infty} c_k p_{x-k} = -\sum_{x=1}^{\infty} b_x \rho^x = 1 - \sum_{x=0}^{\infty} b_x \rho^x =$$

$$= 1 - \exp\left\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k=1}^{\infty} \mathsf{P}\{\zeta_n = k\} \rho^k\right\}$$

(la formule (11) étant vraie pour $|\rho| = 1$, elle le sera pour $|\rho| < 1$).
Donc, eu égard à (16), on trouve

$$(\rho - 1) \sum_{m=0}^{\infty} \rho^m u_m(\lambda, 0) =$$

$$= \exp\left\{\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k=1}^{\infty} \mathsf{P}(\zeta_n = k) \rho^k\right\} \left[\exp\left\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \mathsf{P}\{\zeta_n > 0\}\right\} -\right.$$

$$\left. - \exp\left\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k=1}^{\infty} \mathsf{P}\{\zeta_n = k\} \rho^k\right\}\right] =$$

$$= \exp\left\{\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k=1}^{\infty} \mathsf{P}\{\zeta_n = k\} (\rho^k - 1)\right\} - 1. \quad (21)$$

THÉORÈME 3. *Si* $\mathsf{P}\{\sup_n \zeta_n < \infty\} = 1$, *la répartition de* $\zeta^+ = \sup_n \zeta_n$ *est définie par*

$$\sum_{m=0}^{\infty} \rho^m \mathsf{P}\{\zeta^+ = m\} = \exp\left\{\sum_{n=1}^{\infty} \frac{1}{n} \sum_{k=1}^{\infty} \mathsf{P}\{\zeta_n = k\} (\rho^k - 1)\right\}. \quad (22)$$

Pour le prouver il suffit de passer à la limite pour $\lambda \uparrow 1$ dans (21) et d'appliquer la formule (20). Le théorème 2 autorise ce passage à la limite. ■

REMARQUE. La formule (22) est vraie pour $|\rho| = 1$, car ses deux membres sont définis et continus pour $|\rho| \leqslant 1$.

COROLLAIRE. *Dans les hypothèses du théorème* 3

$$\mathsf{P}\{\sup_{n\geqslant 1} \zeta_n = 0\} =$$

$$= \exp\left\{-\sum_{n=1}^{\infty} \frac{1}{n} \mathsf{P}\{\zeta_n > 0\}\right\}\left[1 - \exp\left\{-\sum_{n=1}^{\infty} \frac{1}{n} \mathsf{P}\{\zeta_n = 0\}\right\}\right]. \quad (23)$$

En effet,

$$\mathsf{P}\{\sup_{n\geqslant 1} \zeta_n = 0\} = \sum_{k=0}^{\infty} p_{-k} \mathsf{P}\{\zeta^+ = k\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \varphi(z) \sum_{k=0}^{\infty} \mathsf{P}\{\zeta^+ = k\} e^{izk}\, dz =$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \varphi(z) \exp\left\{\sum_{n=1}^{\infty} \frac{1}{n} \sum_{k=1}^{\infty} \mathsf{P}\{\zeta_n = k\}(e^{ikz} - 1)\right\} dz =$$

$$= \frac{1}{2\pi} \lim_{\lambda\uparrow 1} \int_{-\pi}^{\pi} \left[\frac{\lambda\varphi(z) - 1}{\lambda} \exp\left\{\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k=1}^{\infty} \mathsf{P}\{\zeta_n = k\} e^{ikz}\right\} + \frac{1}{\lambda}\right] dz \times$$

$$\times \exp\left\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \mathsf{P}\{\zeta_n > 0\}\right\}.$$

En se servant des relations (9) et (7) on peut écrire

$$\mathsf{P}\{\sup_{n\geqslant 1} \zeta_n = 0\} = \frac{1}{2\pi} \lim_{\lambda\uparrow 1} \frac{1}{\lambda} \int_{-\pi}^{\pi} \left[1 - \exp\left\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \times\right.\right.$$

$$\left.\left.\times \sum_{k\leqslant 0} \mathsf{P}\{\zeta_n = k\} e^{ikz}\right\}\right] dz \exp\left\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \mathsf{P}\{\zeta_n > 0\}\right\}. \quad (24)$$

Comme

$$\exp\left\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k\leqslant 0} \mathsf{P}\{\zeta_n = k\} \rho^{-k}\right\}\Big|_{\rho=0} = \exp\left\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \mathsf{P}\{\zeta_n = 0\}\right\},$$

on trouve

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \exp\left\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k\leqslant 0} \mathsf{P}\{\zeta_n = k\} e^{ikz}\right\} dz =$$

$$= \exp\left\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \mathsf{P}\{\zeta_n = 0\}\right\}.$$

En portant cette expression dans (24) on obtient (23).

REMARQUE. *La série*

$$\sum_{n=1}^{\infty} \frac{1}{n} \mathsf{P}\{\zeta_n = 0\} \tag{25}$$

est convergente.

Supposons que $\xi_1 - \xi_2$ admet une répartition réticulée de pas h; alors $|\varphi(z)|^2 = \mathsf{E}e^{iz(\xi_1-\xi_2)}$ est une fonction périodique de période $\frac{\pi}{h}$. Donc

$$\mathsf{P}\{\zeta_n = 0\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi^n(z)\, dz \leqslant \frac{h}{2\pi} \int_{-\pi/h}^{\pi/h} |\varphi(z)|^n\, dz.$$

Le corollaire du lemme 1 dit qu'il existe $\alpha > 0$ tel que $|\varphi(z)|^2 \leqslant e^{-\alpha z^2}$ pour $|z| \leqslant \pi/h$. Par suite

$$\mathsf{P}\{\zeta_n = 0\} \leqslant \frac{h}{2\pi} \int_{-\pi/h}^{\pi/h} e^{-\frac{n\alpha}{2}z^2}\, dz \leqslant \frac{h}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{n\alpha}{2}z^2}\, dz \leqslant \frac{C}{\sqrt{n}},$$

où C est une constante. La convergence de la série (25) résulte de cette majoration.

Plus haut nous avons trouvé seulement la répartition conjointe de τ_0 et γ_0. Pour déterminer celle de τ_x et γ_x on se servira de la relation récurrente suivante:

$$\mathsf{P}\{\tau_x = m,\ \gamma_x = k\} = \mathsf{P}\{\tau_0 = m,\ \gamma_0 = x + k\} + \\ + \sum_{\substack{0<n<m \\ 0<t\leqslant x}} \mathsf{P}\{\tau_0 = n,\ \gamma_0 = t\} \mathsf{P}\{\tau_{x-t} = m - n,\ \gamma_{x-t} = k\}.$$

En multipliant cette expression par λ^m, e^{izk} et en sommant sur m et k de 1 à ∞, on obtient

$$u_x(\lambda,\ z) = \sum_{m,\,k=1}^{\infty} \mathsf{P}\{\tau_0 = m,\ \gamma_0 = x + k\} \lambda^m e^{izk} + \\ + \sum_{0<t\leqslant x} u_{x-t}(\lambda,\ z) \sum_{n=1}^{\infty} \mathsf{P}\{\tau_0 = n,\ \gamma_0 = t\} \lambda^n.$$

Multiplions les deux membres par ρ^x ($|\rho| < 1$) et sommons sur x de 0 à ∞ (pour $x = 0$ la seconde somme du second membre disparaît). On trouve

$$\sum_{x=0}^{\infty} \rho^x u_x(\lambda,\ z) = \sum_{x=0}^{\infty} \rho^x \sum_{m,\,k=1}^{\infty} \mathsf{P}\{\tau_0 = m,\ \gamma_0 = x + k\} \lambda^m e^{izk} + \\ + \sum_{x=0}^{\infty} \rho^x u_x(\lambda,\ z) \sum_{n,\,t=1}^{\infty} \mathsf{P}\{\tau_0 = n,\ \gamma_0 = t\} \lambda^n \rho^t. \tag{26}$$

En vertu de (15)

$$\sum_{n,\,t=1}^{\infty} \lambda^n \rho^t \mathsf{P}\{\tau_0 = n,\ \gamma_0 = t\} = 1 - \exp\Big\{-\sum_{n=1}^{n} \frac{\lambda^n}{n} \sum_{k\geqslant 1} \mathsf{P}\{\zeta_n = k\}\,\rho^k\Big\} \tag{27}$$

(cette formule étant vraie pour $\rho = e^{iz}$, elle le sera pour $|\rho| < 1$). On a ensuite

$$\sum_{\lambda=0}^{\infty} \sum_{m,\,k=1}^{\infty} \mathsf{P}\{\tau_0 = m,\ \gamma_0 = x + k\}\,\lambda^m e^{izk} \rho^x =$$

$$= \sum_{m,\,t=1}^{\infty} \lambda^m \mathsf{P}\{\tau_0 = m,\ \gamma_0 = t\} \sum_{\substack{k+x=t\\ k\geqslant 1,\ x\geqslant 0}} e^{izk} \rho^x =$$

$$= \sum_{m,\,t=1}^{\infty} \lambda^m \mathsf{P}\{\tau_0 = m,\ \gamma_0 = t\}\, \frac{e^{izt} - \rho^t}{1 - \rho e^{-iz}} =$$

$$= \frac{1}{1 - \rho e^{-iz}} \Big[\exp\Big\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k\geqslant 1} \mathsf{P}\{\zeta_n = k\}\,\rho^k\Big\} - \exp\Big\{-\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k\geqslant 1} \mathsf{P}\{\zeta_n = k\}\,e^{izk}\Big\}\Big]$$

(on s'est servi de la formule (27)). Après quelques transformations simples on déduit de (26):

$$\sum_{x=0}^{\infty} \rho^x u_x(\lambda, z) = \frac{1}{1 - \rho e^{-iz}} \Big(1 - \exp\Big\{\sum_{n=1}^{\infty} \frac{\lambda^n}{n} \sum_{k\geqslant 1} \mathsf{P}\{\zeta_n = k\}\,[\rho^k - e^{ikz}]\Big\}\Big). \tag{28}$$

§ 2. Processus discontinu à accroissements indépendants. Processus général de Poisson

Soit $\xi(t)$ $(t \geqslant 0)$ un processus stochastiquement continu à accroissements indépendants à valeurs dans $\mathcal{R}^m$. On dira que ce processus est *discontinu* si pour tout t il existe un nombre fini de points $0 < s_1 < \ldots < s_n \leqslant t$, tels que $\xi(s)$ soit constant sur les intervalles $[0, s_1[$, $]s_1, s_2[$, $\ldots$, $]s_n, t]$. On admettra que le processus est continu à droite: $\xi(s) = \xi(s + 0)$. La continuité stochastique de

$\xi(s)$ entraîne que la probabilité que $\xi(s)$ présente un saut au point s est nulle $\forall\, s \geqslant 0$.

THÉORÈME 1. *Pour qu'un processus stochastiquement continu séparable à accroissements indépendants soit discontinu, il faut et il suffit que pour tout* $t > 0$

$$\overline{\lim_{\lambda_n \to 0}} \sum_{k=1}^{n} \mathsf{P}\{\xi(t_k) - \xi(t_{k-1}) \neq 0\} < \infty,$$

$$0 = t_0 < \ldots < t_n = t, \quad \lambda_n = \max(t_{k+1} - t_k).$$

D é m o n s t r a t i o n. Supposons que le processus est discontinu. Notons $\nu(t)$ le nombre de sauts du processus $\xi(t)$ sur l'intervalle $[0, t]$. Il est évident que $\nu(t)$ est également un processus stochastiquement continu (puisque la probabilité que t soit point de discontinuité de $\nu(t)$ est nulle, car les points de discontinuité de $\xi(t)$ et $\nu(t)$ sont confondus). Par ailleurs, $\nu(t)$ est un processus à accroissements indépendants: $\nu(t+h) - \nu(t)$, le nombre de points de discontinuité de $\xi(s)$ sur $[t, t+h]$, ne dépend pas du comportement de $\xi(s)$ (et partant de $\nu(s)$) sur $[0, t]$. On remarquera que

$$\mathsf{P}\{\xi(t_k) - \xi(t_{k-1}) \neq 0\} \leqslant \mathsf{P}\{\nu(t_k) - \nu(t_{k-1}) \neq 0\}$$

(si $\xi(t_k) \neq \xi(t_{k-1})$, alors $\xi(s)$ présente au moins un saut sur $[t_{k-1}, t_k]$). Donc pour démontrer la condition nécessaire il suffit de prouver que pour tous les $t > 0$

$$\overline{\lim_{\lambda_n \to 0}} \sum_{k=1}^{n} \mathsf{P}\{\nu(t_k) - \nu(t_{k-1}) \neq 0\} < \infty,$$

$$0 = t_0 < \ldots < t_n = t, \quad \lambda_n = \max(t_{k+1} - t_k).$$

Or

$$\exp\left\{-\sum_{k=1}^{n} \mathsf{P}[\nu(t_k) - \nu(t_{k-1}) \neq 0]\right\} \geqslant$$

$$\geqslant \prod_{k=1}^{n} (1 - \mathsf{P}\{\nu(t_k) - \nu(t_{k-1}) \neq 0\}) =$$

$$= \prod_{k=1}^{n} \mathsf{P}\{\nu(t_k) - \nu(t_{k-1}) = 0\} = \mathsf{P}\{\nu(t) = 0\}.$$

La relation $\mathsf{P}\{\nu(t) = 0\} > 0$ résulte de la continuité stochastique uniforme de $\nu(s)$ sur $[0, t]$ et par suite pour $\delta > 0$ on a

$$\mathsf{P}\{\nu(s_1) - \nu(s_2) = 0\} \geqslant 1 - \varepsilon \text{ pour } |s_1 - s_2| < \delta,$$

or

$$\mathsf{P}\{\nu(t) = 0\} = \prod_{k=1}^{n} \mathsf{P}\{\nu(s_k) - \nu(s_{k-1}) = 0\} \geqslant (1 - \varepsilon)^n,$$

où

$$0 = s_0 < s_1 < \ldots < s_n = t \quad \text{et} \quad s_k - s_{k-1} < \delta,$$

donc

$$\sum_{k=1}^{n} \mathsf{P}\{\nu(t_k) - \nu(t_{k-1}) \neq 0\} < -\ln \mathsf{P}\{\nu(t) = 0\} < \infty.$$

D'où la condition nécessaire.

Pour démontrer la condition suffisante on introduira l'instant de premier saut: $\tau_1 = \inf [s: \xi(s) \neq \xi(0)]$ (τ_1 est susceptible de prendre la valeur $+\infty$). Le processus ne présente pas de discontinuités de seconde espèce (cf. chapitre IV, § 4, théorème 3), car séparable. Par suite

$$\mathsf{P}\{\tau_1 > t\} = \lim_{\lambda n \to 0} \mathsf{P}\{\xi(t_1) = \xi(t_0),\ \xi(t_2) = \xi(t_0), \ldots, \xi(t_n) = \xi(t_0)\},$$

$$0 = t_0 < t_1 < \ldots < t_n = t, \quad \lambda_n = \max(t_{k+1} - t_k),$$

$$\mathsf{P}\{\tau_1 > t\} = \lim_{\lambda n \to 0} \prod_{k=1}^{n} (1 - \mathsf{P}\{\xi(t_k) - \xi(t_{k-1}) \neq 0\}).$$

De la continuité stochastique de $\xi(t)$ il résulte que

$$\max_k \mathsf{P}\{\xi(t_{k+1}) - \xi(t_k) \neq 0\} \to 0 \text{ pour } \lambda_n \to 0,$$

donc

$$\prod_{k=1}^{n} (1 - \mathsf{P}\{\xi(t_{k+1}) - \xi(t_k) \neq 0\}) \sim$$

$$\sim \exp\Big\{-\sum_{k=1}^{n} \mathsf{P}\{\xi(t_k) - \xi(t_{k-1}) \neq 0\} \times$$

$$\times (1 + O(\max_k \mathsf{P}\{\xi(t_k) - \xi(t_{k-1}) \neq 0\})\Big\} \sim$$

$$\sim \exp\Big\{-\sum_{k=1}^{n} \mathsf{P}\{\xi(t_k) - \xi(t_{k-1}) \neq 0\}\Big\} \geqslant \alpha > 0.$$

Estimons la probabilité de l'événement: $\xi(s)$ présente r sauts sur $[0, t]$. Soit $p_r(t)$ cette probabilité. On a (exactement comme pour la majoration de $p_0(t)$)

$$p_r(t) = \lim_{\lambda n \to 0} \sum_{1 \leqslant i_1 < i_2 < \ldots < i_r < n} \prod_{j=1}^{r} \mathsf{P}\{\xi(t_{i_j}) - \xi(t_{i_{j-1}}) \neq 0\} \times$$

$$\times \prod_{k \neq i_1, \ldots, i_r} \mathsf{P}\{\xi(t_k) - \xi(t_{k-1}) = 0\},$$

$$0 = t_0 < t_1 < \ldots < t_n = t.$$

Donc

$$p_r(t) = \lim_{\lambda n \to 0} \prod_{k=1}^{n} \mathsf{P}\{\xi(t_k) - \xi(t_{k-1}) = 0\} \times$$
$$\times \sum_{1 \leqslant i_1 < \ldots < i_r \leqslant n} \prod_{j=1}^{r} \mathsf{P}\{\xi(t_{i_j}) - \xi(t_{i_{j-1}}) \neq 0\}$$

(on s'est servi de la relation $\mathsf{P}\{\xi(t_k) - \xi(t_{k-1}) = 0\} \to 1$ uniformément en k pour $\lambda_n \to 0$),

$$\lim_{\lambda n \to 0} \prod_{k=1}^{n} \mathsf{P}\{\xi(t_k) - \xi(t_{k-1}) = 0\} = p_0(t) = \mathsf{P}\{\tau_1 > t\} > 0.$$

D'autre part,

$$\sum_{1 \leqslant i_1 < \ldots < i_r \leqslant n} \prod_{j=1}^{r} \mathsf{P}\{\xi(t_{i_j}) - \xi(t_{i_{j-1}}) \neq 0\} =$$
$$= \frac{1}{r!} \left(\sum_{k=1}^{n} \mathsf{P}\{\xi(t_k) - \xi(t_{k-1}) \neq 0\} \right)^r +$$
$$+ O(\max_j \mathsf{P}\{\xi(t_j) - \xi(t_{j-1}) \neq 0\}) \left(\sum_{k=1}^{n} \mathsf{P}\{\xi(t_k) - \xi(t_{k-1}) \neq 0\} \right)^{r-1}.$$

On a vu plus haut que

$$p_0(t) = \exp\left\{ -\lim_{\lambda n \to 0} \sum_{k=1}^{n} \mathsf{P}\{\xi(t_k) - \xi(t_{k-1}) \neq 0\} \right\}.$$

Si

$$\lambda(t) = -\ln p_0(t),$$

alors

$$p_r(t) = \frac{\lambda^r(t)}{r!} e^{-\lambda(t)}$$

et

$$\sum_{r=0}^{\infty} p_r(t) = \sum_{r=0}^{\infty} \frac{\lambda^r(t)}{r!} e^{-\lambda(t)} = 1.$$

Ceci est la probabilité que le processus $\xi(s)$ admet un nombre fini de sauts sur $[0, t]$. ■

COROLLAIRE. *Si* $\xi(s)$ *est un processus discontinu à accroissements indépendants, alors* $\nu(t)$, *son nombre de sauts sur* $[0, t]$, *est un processus*

poissonien stochastiquement continu à accroissements indépendants

$$\mathsf{P}\{\nu(t)=k\}=\frac{\lambda^k(t)}{k!}e^{-\lambda(t)},$$

où $\lambda(t) = \mathsf{E}\nu(t)$ *est une fonction continue.*

Soit τ_1 la quantité introduite dans la démonstration du théorème 1. Soient

$$\Phi(ds, A) = \mathsf{P}\{\tau_1 \in ds, \ \xi(\tau_1) - \xi(0) \in A\}$$

la répartition conjointe de τ_1 et $\xi(\tau_1) - \xi(0)$, $\pi(s, A)$ la répartition conditionnelle de $\xi(\tau_1) - \xi(0)$ sachant que $\tau_1 = s$:

$$\int\limits_{\alpha\leqslant s\leqslant\beta} \pi(s, A)\,\mathsf{P}\{\tau_1 \in ds\} = \\ = \mathsf{P}\{\tau_1 \in [\alpha, \beta], \ \xi(\tau_1) - \xi(0) \in A\} = \\ = \mathsf{P}\{\xi(u) = \xi(0), \ u \leqslant \alpha, \ \xi(\tau^\alpha) - \xi(\alpha) \in A, \ \tau^\alpha < \beta\},$$

où $\tau^\alpha = \inf[s > \alpha : \xi(s) - \xi(\alpha) \neq 0]$. Comme

$$\mathsf{P}\{\tau_1 < s\} = 1 - \exp\{-\lambda(s)\},$$

la dernière égalité s'écrit:

$$\int\limits_{\alpha\leqslant s\leqslant\beta} \pi(s, A)\,e^{-\lambda(s)}d\lambda(s) = e^{-\lambda(\alpha)}\mathsf{P}\{\xi(\tau^\alpha)-\xi(\alpha)\in A, \ \tau^\alpha<\beta\},$$

d'où

$$\mathsf{P}\{\xi(\tau^\alpha)-\xi(\alpha)\in A, \ \tau^\alpha<\beta\}=\int\limits_{\alpha\leqslant s\leqslant\beta}\pi(s, A)\,e^{-[\lambda(s)-\lambda(\alpha)]}d\lambda(s).$$

Cette formule permet de déterminer la fonction caractéristique d'un processus à accroissements indépendants. On remarquera que pour $z \in \mathscr{R}^n$

$$\mathsf{E}e^{i(z,\,\xi(\beta)-\xi(\alpha))} = \mathsf{E}\chi_{\{\tau^\alpha>\beta\}} + \\ + \mathsf{E}e^{i(z,\,\xi(\beta)-\xi(\alpha))}\chi_{\{\nu(\beta)-\nu(\alpha)=1\}} + O(\mathsf{P}\{\nu(\beta)-\nu(\alpha)>1\}).$$

Or, si $\nu(\beta)-\nu(\alpha)=1$, on a $\xi(\beta)-\xi(\alpha)=\xi(\tau^\alpha)-\xi(\alpha)$ et $\tau^\alpha<\beta$. Donc

$$\mathsf{E}\chi_{\{\tau^\alpha>\beta\}} + \mathsf{E}e^{i(z,\,\xi(\beta)-\xi(\alpha))}\chi_{\{\nu(\beta)-\nu(\alpha)=1\}} = \\ = 1 + \mathsf{E}\left[e^{i(z,\,\xi(\tau^\alpha)-\xi(\alpha))} - 1\right]\chi_{\{\tau^\alpha<\beta\}} + \\ + O(\mathsf{P}\{\nu(\beta)-\nu(\alpha)>1\}) = 1 + \int\limits_{\alpha\leqslant s\leqslant\beta}\int (e^{i(z,\,x)}-1)\,\pi(s, dx)\times \\ \times e^{-[\lambda(s)-\lambda(\alpha)]}d\lambda(s) + O([\lambda(\beta)-\lambda(\alpha)]^2) = \\ = \exp\Big\{\int\limits_{\alpha\leqslant s\leqslant\beta}\int (e^{i(z,\,x)}-1)\,\pi(s, dx)\,d\lambda(s) + O([\lambda(\beta)-\lambda(\alpha)]^2)\Big\}.$$

Par suite, pour $u = t_0 < t_1 < \ldots < t_n = t$

$$\mathsf{E}\exp\{i(z, \xi(t)-\xi(u))\} = \\ = \exp\Big\{\sum_{k=1}^{n} \int_{t_{k-1}<t<t_k} \int (e^{i(z,x)}-1)\,\pi(s, dx)\,d\lambda(s) + \\ + O\Big(\sum_{k=1}^{n} [\lambda(t_k)-\lambda(t_{k-1})]^2\Big)\Big\}.$$

En passant à la limite pour $\max(t_k - t_{k-1}) \to 0$ on a le

THÉORÈME 2. *Si $\xi(t)$ est un processus discontinu à accroissements indépendants, son accroissement $\xi(t) - \xi(u)$ $(u < t)$ admet la fonction caractéristique :*

$$\mathsf{E}\exp\{i(z, \xi(t)-\xi(u))\} = \exp\Big\{\int_u^t \int_{\mathcal{R}^m} (e^{i(z,x)}-1)\,\pi_s(dx)\,d\lambda(s)\Big\}, \quad (1)$$

où $e^{-\lambda(s)} = \mathsf{P}\{\tau_1 > s\}$ est la répartition de l'instant de premier saut et $\pi_s(A)$, la répartition conditionnelle du premier saut sachant qu'il s'est produit à l'instant s.

REMARQUE. De (1) il suit que la fonction caractéristique $\xi(t) - \xi(0)$ s'écrit

$$\exp\Big\{\int_{\mathcal{R}^m} (e^{i(z,x)}-1)\,\Pi_t(dx)\Big\}, \quad (2)$$

où $\Pi_t(A) = \int_0^t \pi_s(A)\,d\lambda(s)$ est une mesure finie sur $\mathcal{R}^m$. Montrons que cette condition est suffisante pour que le processus soit discontinu. En effet, si ξ admet une répartition indéfiniment divisible, de fonction caractéristique

$$e^{-\Pi(\mathcal{R}^m)}\exp\Big\{\Pi(\mathcal{R}^m)\int e^{i(z,x)}\frac{\Pi(dx)}{\Pi(\mathcal{R}^m)}\Big\} = \\ = \sum_{n=0}^{\infty} \frac{\Pi^n(\mathcal{R}^m)}{n!} e^{-\Pi(\mathcal{R}^m)} (\varphi(z))^n,$$

$$\varphi(z) = \frac{1}{\Pi(\mathcal{R}^m)} \int e^{i(z,x)}\Pi(dx),$$

elle admettra la même répartition que $\sum_1^{\nu} \eta_k$, où η_k sont indépendants et de fonction caractéristique $\varphi(z)$, et ν possède une répartition de

Poisson de paramètre $\Pi(\mathcal{R}^m)$. Donc

$$\mathsf{P}\{\xi \neq 0\} \leqslant \mathsf{P}\{\nu \neq 0\} \leqslant 1 - e^{-\Pi(\mathcal{R}^m)}.$$

Donc si $\xi(t) - \xi(0)$ possède la fonction caractéristique (2), alors

$$\mathsf{P}\{\xi(t_k) - \xi(t_{k-1}) \neq 0\} \leqslant$$
$$\leqslant 1 - \exp\{-\Pi_{t_k}(\mathcal{R}^m) + \Pi_{t_{k-1}}(\mathcal{R}^m)\} \leqslant \Pi_{t_k}(\mathcal{R}^m) - \Pi_{t_{k-1}}(\mathcal{R}^m).$$

D'où la condition du théorème 1.

Soit maintenant $\xi(t)$ processus discontinu homogène (cf. chapitre I, § 3). Le processus $\nu(t)$ sera visiblement homogène. Donc la fonction $\lambda(t)$ est de la forme λt, où λ est une constante. De (1) il résulte que

$$\mathsf{E}\exp\{i(z, \xi(t+h) - \xi(t))\} =$$
$$= \exp\left\{\lambda \int_t^{t+h} \int_{\mathcal{R}^m} (e^{i(z,x)} - 1)\, \pi_s(dx)\, ds\right\} =$$
$$= \exp\left\{\lambda \int_0^h \int_{\mathcal{R}^m} (e^{i(z,x)} - 1)\, \pi_{s+t}(dx)\, ds\right\},$$

et la dernière expression ne doit pas dépendre de t. D'où il suit que $\pi_s(A)$ ne dépend pas de s. Par suite $\xi(\tau_1) - \xi(0)$ et τ_1 sont indépendants.

THÉORÈME 3. *Si* $\xi(t)$ $(\xi(0) = 0)$ *est un processus homogène discontinu à accroissements indépendants, le processus* $\xi_1(t) = \xi(t + \tau_1) - \xi(\tau_1)$ *est aussi un processus homogène à accroissements indépendants ne dépendant pas de* τ_1 *et* $\eta_1 = \xi(\tau_1)$.

D é m o n s t r a t i o n. Soit $0 < s_1 < \ldots < s_k$, $z_0, z_1, \ldots, z_k \in \mathcal{R}^m$. On a

$$\mathsf{E}e^{-\lambda\tau_1 + i(z_0, \eta_1)} \exp\left\{i \sum (z_j, \xi(s_j + \tau_1) - \xi(\tau_1))\right\} =$$
$$= \lim_{h\to 0} \mathsf{E}\exp\left\{-\lambda\tau^{(h)} + i(z_0, \eta_1^{(h)}) + i \sum_{j=1}^k (z_j, \xi(s_j + \tau^{(h)}) - \xi(\tau^{(h)}))\right\}, \tag{3}$$

où $\tau^{(h)} = nh$ si $\xi(lh) = 0$ pour $l = 1, \ldots, n-1$, $\xi(nh) \neq 0$.

D'autre part

$$\mathsf{E}\exp\left\{-\lambda\tau^{(h)} + i(z_0, \eta_1^{(h)}) + i \sum_{j=1}^k (z_j, \xi(s_j + \tau^{(h)}) - \xi(\tau^{(h)}))\right\} =$$
$$= \sum_{n=1}^{\infty} \mathsf{E}\exp\left\{-\lambda nh + i(z_0, \xi(nh)) + \right.$$

$$+i \sum_{j=1}^{k} (z_j, \xi(s_j+nh)-\xi(nh)) \Big\} \chi_{\{\tau^{(h)}=nh\}} =$$

$$= \mathsf{E} \exp\Big\{ i \sum_{j=1}^{k} (z_j, \xi(s_j)) \Big\} \sum_{n=1}^{\infty} \mathsf{E} \exp\{-\lambda nh +$$

$$+ i(z_0, \xi(nh))\} \chi_{\{\tau^{(h)}=nh\}} =$$

$$= \mathsf{E} e^{-\lambda\tau^{(h)}+i(z_0, \xi(nh))} \mathsf{E} \exp\Big\{ i \sum_{j+1}^{k} (z_j, \xi(s_j)) \Big\}.$$

Donc, en passant à la limite dans le second membre de (3), on trouve

$$\mathsf{E} e^{-\lambda\tau_1+i(z_0, \eta_1)} \exp\Big\{ i \sum_{j=1}^{k} (z_j, \xi(s_j+\tau_1)-\xi(\tau_1)) \Big\} =$$

$$= \mathsf{E} e^{-\lambda\tau_1+i(z_0, \eta_1)} \mathsf{E} \exp\Big\{ i \sum_{j=1}^{k} (z_j, \xi(s_j)) \Big\}. \blacksquare$$

COROLLAIRE. *Soient* $\eta_1, \eta_2, \ldots$ *les sauts consécutifs du processus* $\xi(t)$, $\tau_1, \tau_2, \ldots$ *les intervalles séparant les instants de ces sauts; toutes ces variables aléatoires sont indépendantes dans leur ensemble.*

En effet, $\eta_2, \eta_3, \ldots, \tau_2, \tau_3, \ldots$ étant entièrement déterminés par le processus $\xi_1(t)$, les variables η_1 et τ_1 ne dépendent pas d'eux et de plus sont mutuellement indépendantes.

Comme $\xi_1(t)$ s'exprime de façon analogue en fonction de $\eta_2, \eta_3, \ldots, \tau_2, \tau_3, \ldots$, les variables η_2 et τ_2 sont indépendantes et ne dépendent pas de $\eta_3, \ldots, \tau_3, \ldots$ En envisageant de proche en proche les processus $\xi_n(t) = \xi_{n-1}(t+\tau_n) - \xi_{n-1}(\tau_n)$, on s'assure que les n variables η_n et τ_n sont indépendantes et ne dépendent pas de $\eta_{n+1}, \tau_{n+1}, \ldots$

Soit $\nu(t)$ le nombre de sauts du processus $\xi(s)$ sur l'intervalle $[0, t]$. Comme

$$\nu(t) = m \quad \text{si} \quad \sum_{1}^{m} \tau_k \leqslant t < \sum_{1}^{m+1} \tau_k \ \Big(\sum_{1}^{0} = 0\Big),$$

le processus $\nu(t)$ s'exprime en fonction de τ_k, $k = 1, \ldots$, et par suite ne dépend pas de η_k, $k = 1, \ldots$

Donc

$$\xi(t) = \sum_{k=1}^{\nu(t)} \eta_k, \tag{4}$$

où $\eta_1, \ldots, \eta_k$ est une suite de variables indépendantes équi-réparties et $\nu(t)$ un processus poissonien homogène ne dépendant pas d'elles. Le processus (4) est dit *processus poissonien général.*

Etudions la répartition de certaines caractéristiques d'un tel processus dans $\mathscr{R}^1$. Soient a paramètre d'un processus poissonien, $\Phi(x)$ fonction de répartition des variables η_k. Trouvons la fonction

de répartition du maximum du processus sur un intervalle fini. Soit

$$Q(t, x) = \mathsf{P}\{\sup_{s\leqslant t} \xi(s) < x\}.$$

L'indépendance de τ_1, η_1 et du processus $\xi_1(t)$ nous permet d'écrire pour $x > 0$

$$Q(t, x) = \mathsf{P}\{\tau_1 > t\} + \mathsf{P}\{\sup_{0\leqslant s\leqslant t} \xi(s) < x,\ \tau_1 \leqslant t\} =$$
$$= \mathsf{P}\{\tau_1 > t\} + \mathsf{P}\{\tau_1 \leqslant t,\ \eta_1 < x,\ \sup_{0<s<t-\tau_1} \xi_1(s) < x - \eta_1\} =$$
$$= e^{-at} + \int_0^t \int_{-\infty}^x \mathsf{P}\{\tau_1 \in du,\ \eta_1 \in dy\}\, \mathsf{P}\{\sup_{0<s<t-u} \xi_1(s) < x-y\} =$$
$$= e^{-at} + \int_0^t e^{-au}\, du \int_{-\infty}^x d\Phi(y)\, Q(t-u, x-y).$$

Pour $x \leqslant 0$ $Q(t, x) = 0$. Donc en posant

$$\int_0^\infty e^{-\lambda t} Q(t, x)\, dt = q(\lambda, x),$$

on trouve

$$q(\lambda, x) = \frac{1}{a+\lambda} + \frac{a}{a+\lambda} \int q(\lambda, x-y)\, d\Phi(y) \tag{5}$$

($q(\lambda, x) = 0$ pour $x \leqslant 0$). La relation (5) est vraie pour tous les $x > 0$. Donc si l'on pose

$$\varepsilon(x) = \begin{cases} 1 & \text{pour } x > 0, \\ 0 & \text{pour } x \leqslant 0, \end{cases}$$

on peut mettre (5) sous la forme

$$\frac{\varepsilon(x)}{a+\lambda} = \varepsilon(x) \int q(\lambda, x-y)\, d\left[\varepsilon(y) - \frac{a}{a+\lambda}\Phi(y)\right]. \tag{6}$$

Nous allons montrer comment on peut résoudre l'équation

$$g(x) = \varepsilon(x) \int q(x-y)\, d\left[\varepsilon(y) - \frac{a}{a+\lambda}\Phi(y)\right], \tag{7}$$
$$q(x) = 0,\ x \leqslant 0,$$

où $g(x)$ est une fonction bornée, $g(x) = 0$ pour $x \leqslant 0$. Soit $v_1(t)$ une fonction à variation bornée telle que $v_1(t) = v_1(0)$ pour $t > 0$. De (7) il suit

$$\int g(x-t)\, dv_1(t) =$$
$$= \int \varepsilon(x-t) \int q(x-t-y)\, d\left[\varepsilon(y) - \frac{a}{a+\lambda}\Phi(y)\right] dv_1(t).$$

Si $x > 0$, on peut substituer $\varepsilon(x)$ à $\varepsilon(x-t)$ dans le second membre, puisque l'intégration est effectuée sur les seuls t négatifs. Donc

$$\varepsilon(x)\int g(x-t)\,dv_1(t) =$$
$$= \varepsilon(x)\int\int q(x-t-y)\,d\left[\varepsilon(y)-\frac{a}{a+\lambda}\Phi(y)\right]dv_1(t) =$$
$$= \varepsilon(x)\int q(x-y)\,dv_2(y),$$

où la fonction à variation bornée $v_2(y)$ est définie par

$$v_2(y) = \int v_1(y-t)\,d\left[\varepsilon(t)-\frac{a}{a+\lambda}\Phi(t)\right]. \qquad (8)$$

Soit maintenant $v_2(y)$ tel que $v_2(y)=0$ pour $y \leqslant 0$. On a

$$\varepsilon(x)\int q(x-y)\,dv_2(y) = \int q(x-y)\,dv_2(y),$$

puisque pour $x \leqslant 0$

$$\int q(x-y)\,dv_2(y) = \int_0^\infty q(x-y)\,dv_2(y) = 0.$$

Donc, dans le cas où existent des fonctions $v_1(x)$ et $v_2(x)$ possédant les propriétés indiquées, on déduit de (7)

$$\varepsilon(x)\int g(x-y)\,dv_1(y) = \int q(x-y)\,dv_2(y). \qquad (9)$$

La dernière équation en q est une équation de convolution soluble par la transformation de Fourier. La méthode de résolution de l'équation (7) (appelée équation de convolution sur le demi-axe) proposée ici est due à N. Wiener.

Voyons comment on trouve des fonctions $v_1(x)$ et $v_2(x)$ possédant les propriétés requises. Posons

$$\tilde{v}_k(z) = \int e^{izx}\,dv_k(x), \quad \varphi(z) = \int e^{izx}\,d\Phi(x).$$

En appliquant la transformation de Fourier à (8), on obtient

$$\tilde{v}_2(z) = \tilde{v}_1(z)\left[1-\frac{a}{a+\lambda}\varphi(z)\right]. \qquad (10)$$

Mais

$$1-\frac{a}{a+\lambda}\varphi(z) = \exp\left\{\ln\left(1-\frac{a}{a+\lambda}\varphi(z)\right)\right\} =$$
$$= \exp\left\{-\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{a}{a+\lambda}\right)^n\varphi^n(z)\right\}.$$

On a $\varphi^n(z) = \int e^{izx}\, d\Phi_n(x)$, où $\Phi_n(x)$ est fonction de répartition de $\eta_1 + \ldots + \eta_n$.

Soit

$$\widetilde{v}_1(z) = \exp\left\{\sum_{n=1}^{\infty} \frac{1}{n}\left(\frac{a}{a+\lambda}\right)^n \int_{-\infty}^{+0} e^{izx}\, d\Phi_n(x)\right\}, \tag{11}$$

$$\widetilde{v}_2(z) = \exp\left\{-\sum_{n=1}^{\infty} \frac{1}{n}\left(\frac{a}{a+\lambda}\right)^n \int_{+0}^{\infty} e^{izx}\, d\Phi_n(x)\right\}. \tag{12}$$

La relation (10) est réalisée. Soit

$$H_-(x) = \begin{cases} \sum_{n=1}^{\infty} \frac{1}{n}\left(\frac{a}{a+\lambda}\right)^n \Phi_n(x), & x < 0, \\ \sum_{n=1}^{\infty} \frac{1}{n}\left(\frac{a}{a+\lambda}\right)^n \Phi_n(0), & x > 0. \end{cases}$$

Alors

$$\widetilde{v}_1(z) = 1 + \sum_{k=1}^{\infty} \frac{1}{k!}\left[\int e^{izx}\, dH_-(x)\right]^k =$$

$$= \int e^{izx}\, d\left[\varepsilon(x) + \sum_{k=1}^{\infty} \frac{1}{k!} H_-^{(k)}(x)\right],$$

où $H_-^{(1)}(x) = H_-(x)$, $H_-^{(k)}(x) = \int H_-^{(k-1)}(x-y)\, dH_-(y)$. Il est manifeste que la fonction

$$v_1(x) = \varepsilon(x) + \sum_{k=1}^{\infty} \frac{1}{k!} H_-^{(k)}(x)$$

est à variation bornée et que $v_1(x) = v_1(0)$ pour $x \geqslant 0$. De façon analogue on s'assure que

$$\widetilde{v}_2(z) = \int e^{izx}\, d\left[\varepsilon(x) + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} H_+^{(k)}(x)\right],$$

où

$$H_+(x) = \begin{cases} 0, & x \leqslant 0, \\ \sum_{n=1}^{\infty} \frac{1}{n}\left(\frac{a}{a+\lambda}\right)^n \left[\Phi_n(x) - \Phi_n(0)\right], & x > 0, \end{cases}$$

$$H_+^{(1)}(x)=H_+(x),\quad H_+^{(k)}(x)=\int H_+^{(k-1)}(x-y)\,dH_+(y).$$

La fonction

$$v_2(x)=\varepsilon(x)+\sum_{k=1}^{\infty}\frac{(-1)^k}{k!}H_+^{(k)}(x)$$

est également à variation bornée et $v_2(x)=0$ pour $x\leqslant 0$.

Déterminons maintenant la fonction $q(\lambda, x)$. De (6), (7) et (9) il résulte que

$$\varepsilon(x)\int\frac{\varepsilon(x-1)}{a+\lambda}dv_1(y)=\int q(\lambda, x-y)\,dv_2(y). \tag{13}$$

Le premier membre étant confondu avec $\frac{v_1(0)}{a+\lambda}$ pour $x>0$, (13) équivaut à

$$\frac{\varepsilon(x)\,v_1(0)}{a+\lambda}=\int q(\lambda, x-y)\,dv_2(y).$$

La transformation de Fourier donne

$$\frac{v_1(0)}{a+\lambda}=\widetilde{v}_2(z)\int e^{izx}\,dq(\lambda, x).$$

Comme $v_1(0)=v_1(+\infty)=\widetilde{v}_1(0)$,

$$\frac{1}{a+\lambda}=\frac{1}{\lambda}\left(1-\frac{a}{a+\lambda}\right)=$$

$$=\frac{1}{\lambda}\exp\left\{\ln\left(1-\frac{a}{a+\lambda}\right)\right\}=\frac{1}{\lambda}\exp\left\{-\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{a}{a+\lambda}\right)^n\right\},$$

on trouve

$$\int e^{izx}\,dq(\lambda, x)=\frac{1}{\lambda}\frac{\widetilde{v}_1(0)}{\widetilde{v}_2(z)}\exp\left\{-\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{a}{a+\lambda}\right)^n\right\}=$$

$$=\frac{1}{\lambda}\exp\left\{\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{a}{a+\lambda}\right)^n\left[\int_{-\infty}^{+0}d\Phi_n(x)+\int_{+0}^{\infty}e^{izx}\,d\Phi_n(x)-1\right]\right\}.$$

Et en définitive

$$\int e^{izx}\,dq(\lambda, x)=\frac{1}{\lambda}\exp\left\{\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{a}{a+\lambda}\right)^n\int_0^{\infty}(e^{izx}-1)\,d\Phi_n(x)\right\}. \tag{14}$$

La représentation (4) nous permet d'écrire

$$\mathsf{P}\{\xi(t)<x\}=\varepsilon(x)+\sum_{n=1}^{\infty}\frac{(at)^n}{n!}e^{-at}\Phi_n(x). \tag{15}$$

On remarquera maintenant que

$$\frac{1}{n}\left(\frac{a}{a+\lambda}\right)^n = \int_0^\infty e^{-\lambda t}\frac{a^n t^{n-1}}{n!}e^{-at}\,dt.$$

Donc, en notant $F_t(x) = \mathsf{P}\{\xi(t) < x\}$, on aura $\left(\text{puisque } \int_0^\infty (e^{izx}-1)\,d\varepsilon(x) = 0\right)$

$$\int_0^\infty \frac{e^{-\lambda t}}{t}\int_0^\infty (e^{izx}-1)\,dF_t(x)\,dt =$$

$$= \sum_{n=1}^\infty \int_0^\infty e^{-\lambda t}\frac{a^n t^{n-1}}{n!}e^{-at}\int_0^\infty (e^{izx}-1)\,d\Phi_n(x)\,dt =$$

$$= \sum_{n=1}^\infty \frac{1}{n}\left(\frac{a}{a+\lambda}\right)^n \int_0^\infty (e^{izx}-1)\,d\Phi_n(x).$$

On peut donc exprimer le second membre de (14) directement en fonction de la répartition de $\xi(t)$.

Théorème 4. *Soit $\xi(t)$ un processus poissonien général, $F_t(x) = \mathsf{P}\{\xi(t) < x\}$. La fonction*

$$\tilde{q}(\lambda, z) = \int_0^\infty \int e^{-\lambda t + izx}\,d\mathsf{P}\{\sup_{s\leqslant t}\xi(s) < x\}\,dt$$

admet alors la représentation

$$\tilde{q}(\lambda, z) = \frac{1}{\lambda}\exp\left\{\int_0^\infty \frac{e^{-\lambda t}}{t}\int_0^\infty (e^{izx}-1)\,dF_t(x)\,dt\right\}. \tag{16}$$

Corollaire. *Pour que*

$$\mathsf{P}\{\sup_{t\geqslant 0}\xi(t) < +\infty\} = 1,$$

il est nécessaire et suffisant que

$$\int_0^\infty \frac{1}{t}\mathsf{P}\{\xi(t) > 0\}\,dt < \infty. \tag{17}$$

Si cette condition est réalisée, la répartition de la variable $\xi_+ = \sup_{t\geq 0} \xi(t)$ *est définie par la fonction caractéristique*:

$$\mathsf{E} e^{iz\xi_+} = \exp\left\{\int_0^\infty \frac{1}{t} \int_0^\infty (e^{izx}-1)\, dF_t(x)\right\}. \tag{18}$$

En effet,

$$\mathsf{P}\{\sup_{t\geq 0} \xi(t) < +\infty\} = \mathsf{P}\left\{\sup_n \sum_{k=1}^n \eta_k = +\infty\right\}.$$

Le théorème 2, § 1, dit que la dernière probabilité est égale à 1 si et seulement si

$$\sum_{n=1}^\infty \frac{1}{n}(1-\Phi_n(0)) < \infty.$$

La relation (15) nous montre que

$$\sum_{n=1}^\infty \frac{1}{n}(1-\Phi_n(0)) = \int_0^\infty \frac{1}{t}(1-F_t(0))\, dt.$$

La formule (18) se déduit de (16) par un passage à la limite dont la légitimité résulte de la finitude de la mesure

$$\int_0^\infty \frac{1}{t} \int_A dF_t(x)\, dt$$

sur $[0, \infty[$, laquelle découle de (17).

Soit $x \geqslant 0$. Posons

$$\tau^x = \inf[t : \xi(t) > x], \quad \gamma_x = \xi(\tau^x+0) - \xi(\tau^x);$$

τ^x est l'*instant de premier saut du niveau* x, γ_x *la valeur du premier saut du niveau* x. Si $\sup_{s<\infty} \xi(s) \leqslant x$, on pose $\tau^x = +\infty$; dans ce cas γ_x n'est pas défini.

Cherchons la répartition conjointe de τ^x et γ_x. Désignons

$$N(t, y, x) = \mathsf{P}\{\tau^x < t, \gamma_x \geqslant y\}.$$

De toute évidence, si $\eta_1 > x$, $\tau^x = \tau_1$, $\gamma_x = \eta_1 - x$. Si $\eta_1 \leqslant x$, alors

$$\tau^x = \hat{\tau}^{x-\xi_1} + \tau_1, \quad \gamma_x = \hat{\gamma}_{x-\xi_1},$$

où $\hat{\tau}^y$, $\hat{\gamma}_y$ sont respectivement l'instant et la valeur du saut du processus $\xi_1(t) = \xi(t+\tau_1) - \xi(\tau_1)$. Le processus $\xi_1(t)$ étant indé-

pendant de τ_1 et η_1, on obtient l'équation suivante:

$$N(t, y, x) = \mathsf{P}\{\tau_1 < t\}\,\mathsf{P}\{\eta_1 \geqslant x+y\} + \int_0^t \int_{-\infty}^x a e^{-as} N(t-s, y, x-u)\, d\Phi(u)\, ds. \qquad (19)$$

Soit

$$n(\lambda, y, x) = \int_0^\infty e^{-\lambda t}\, d_t N(t, y, x).$$

En appliquant la transformation de Laplace à (19) on trouve

$$n(\lambda, y, x) = \frac{a}{a+\lambda}[1-\Phi(x+y)] + \frac{a}{a+\lambda}\int_{-\infty}^x n(\lambda, y, x-u)\, d\Phi(u). \qquad (20)$$

Etudions cette équation pour y fixe. On admettra que $n(\lambda, y, x) = 0$ pour $x < 0$. Cette équation s'écrit alors

$$\varepsilon(x)\frac{a}{a+\lambda}[1-\Phi(x+y)] = \varepsilon(x)\int_{-\infty}^\infty n(\lambda, y, x-u)\, d\left[\varepsilon(u) - \frac{a}{a+\lambda}\Phi(u)\right]. \qquad (21)$$

C'est une équation de la forme (7). Donc elle est justiciable de la relation (9):

$$\varepsilon(x)\int \varepsilon(x-u)\frac{a}{a+\lambda}[1-\Phi(x-u+y)]\, dv_1(u) = \int n(\lambda, y, x-u)\, dv_2(u).$$

En multipliant par $e^{-\mu x}$ et en intégrant sur x de 0 à ∞, on trouve

$$\int_0^\infty \int_0^\infty \varepsilon(x-u)\frac{a}{a+\lambda}[1-\Phi(x-u+y)]\, e^{-\mu x}\, dv_1(u)\, dx = \int_0^\infty e^{-\mu u}\, dv_2(u) \int_0^\infty n(\lambda, y, x)\, e^{-\mu x}\, dx.$$

De (12) il résulte que

$$\int_0^\infty e^{-\mu u}\, dv_2(u) = \exp\left\{-\sum_{n=1}^\infty \frac{1}{n}\left(\frac{a}{a+\lambda}\right)^n \int_0^\infty e^{-\mu u}\, d\Phi_n(u)\right\}.$$

Donc

$$\int_0^\infty n(\lambda, y, x) e^{-\mu x} dx = \exp\left\{\sum_{n=1}^\infty \frac{1}{n}\left(\frac{a}{a+\lambda}\right)^n \int_0^\infty e^{-\mu u} d\Phi_n(u)\right\} \times$$

$$\times \int_0^\infty e^{-\mu x} \int_0^\infty \frac{a}{a+\lambda}[1-\Phi(x-u+y)]\, dv_1(u)\, dx.$$

Le fait que la transformée de Laplace du produit de convolution de deux fonctions est égale au produit des transformées de Laplace, et l'égalité

$$\exp\left\{\sum_{n=1}^\infty \frac{1}{n}\left(\frac{a}{a+\lambda}\right)^n \int_0^\infty (e^{-\mu u}-1)\, d\Phi_n(u)\right\} = \lambda \int_0^\infty e^{-\mu x} dq(\lambda, x),$$

qui découle de (14), donnent

$$n(\lambda, y, x) = \lambda \exp\left\{\sum_{n=1}^\infty \frac{1}{n}\left(\frac{a}{a+\lambda}\right)^n \int_0^\infty d\Phi_n(u)\right\} \times$$

$$\times \int\int \frac{a}{a+\lambda}[1-\Phi(x-z-u+y)]\, dv_1(u)\, dq(\lambda, z). \quad (22)$$

Supposons que $q_-(\lambda, x)$ se tire de l'équation

$$\int e^{izx} dq_-(\lambda, x) = \frac{1}{\lambda} \exp\left\{\sum_{n=1}^\infty \frac{1}{n}\left(\frac{a}{a+\lambda}\right)^n \int_{-\infty}^0 (e^{izx}-1)\, d\Phi_n(x)\right\}. \quad (23)$$

On s'assure immédiatement en envisageant le processus $-\xi(t)$ que

$$q_-(\lambda, x) = \int_0^\infty e^{-\lambda t} \mathsf{P}\{\inf_{s \leqslant t} \xi(s) < x\}\, dt. \quad (24)$$

En vertu de (11) la fonction $q_-(\lambda, x)$ s'exprime en fonction de $v_1(x)$ comme suit:

$$q_-(\lambda, x) = \frac{1}{\lambda} \exp\left\{-\sum_{n=1}^\infty \frac{1}{n}\left(\frac{a}{a+\lambda}\right)^n \int_{-\infty}^0 d\Phi_n(x)\right\} v_1(x).$$

En portant dans (22) l'expression de $v_1(x)$ en fonction de $q_-(\lambda, x)$ et compte tenu de

$$\exp\left\{\sum_{n=1}^\infty \frac{1}{n}\left(\frac{a}{a+\lambda}\right)^n\right\} = \exp\left\{-\ln\left(1-\frac{a}{a+\lambda}\right)\right\} = \frac{1}{1-\frac{a}{a+\lambda}} = \frac{a+\lambda}{\lambda},$$

on trouve finalement

$$n(\lambda, y, x) = \lambda a \int\int [1-\Phi(x+y-u-z)]\, dq_-(\lambda, z)\, dq(\lambda, u). \quad (25)$$

Inversons (25) en utilisant encore le fait que le produit de transformées de Laplace est la transformée de Laplace d'un produit de convolution. A remarquer que

$$\lambda q_-(\lambda, z) = -\int_0^\infty Q_-(t, x)\, de^{-\lambda t} =$$

$$= Q_-(0, x) + \int_0^\infty Q_-(t, x)\, e^{-\lambda t}\, dt = \varepsilon(x) + \int_0^\infty e^{-\lambda t} d_t Q_-(t, x),$$

où

$$Q_-(t, x) = \mathsf{P}\{\inf_{s \leqslant t} \xi(s) < x\}.$$

Donc $\int \lambda q_-(\lambda, z-u)\, dq(\lambda, u)$ est la transformée de Laplace de la fonction

$$Q(t, z) + \int_0^t \int Q_-(t-s, z-u)\, d_u d_s Q(s, u).$$

Par suite

$$N(t, y, x) = a \int [1 - \Phi(x+y-z)]\, d_z Q(t, z) +$$

$$+ a \int [1 - \Phi(x+y-z)]\, d_z \int_0^t \int Q_-(t-s, z-u)\, d_u d_s Q(s, u). \quad (26)$$

§ 3. Processus continus. Processus wienérien

Dans ce paragraphe on se propose d'étudier un processus continu à accroissements indépendants $\xi(t)$, défini sur un intervalle $[0, T]$ et à valeurs dans $\mathscr{R}^m$. On montrera que les accroissements du processus admettent des répartitions gaussiennes. Appelons $\mathscr{L}_+(\mathscr{R}^m)$ l'ensemble des opérateurs symétriques linéaires non négatifs de $\mathscr{R}^m$.

Théorème 1. *Il existe des fonctions continues* $a(t)$ *et* $B(t)$ *à valeurs dans* $\mathscr{R}^m$ *et* $\mathscr{L}_+(\mathscr{R}^m)$ *respectivement* ($B(t)$ *ne décroît pas,* $B(t_2) - B_1(t_1) \in \mathscr{L}_+(\mathscr{R}^m)$ *pour* $t_1 < t_2$) *telles que la fonction caractéristique de l'accroissement* $\xi(t_2) - \xi(t_1)$ $(t_1 < t_2)$ *est de la forme*

$$\mathsf{E} \exp\{i(z, \xi(t_2) - \xi(t_1))\} =$$

$$= \exp\left\{i(z, a(t_2) - a(t_1)) - \frac{1}{2}([B(t_2) - B(t_1)]\, z, z)\right\}. \quad (1)$$

Démonstration. Pour établir (1) il suffit de prouver que $\xi(t_2) - \xi(t_1)$ admet une répartition normale. On suppose que

$\xi(0) = 0$, $t_1 = 0$ et que le processus prend ses valeurs dans $\mathscr{R}^1$ (au lieu de $\xi(t)$ on peut considérer le processus $(\xi(t), z)$). Prenons une suite de subdivisions $0 = t_{n0} < \ldots < t_{nn} = t$ telle que $\max_k (t_{nk} - t_{n,k-1}) \to 0$. La continuité du processus implique

$$\lim_{n\to\infty} \sum_{k=1}^{n} \mathsf{P}\{|\xi(t_{nk}) - \xi(t_{n,k-1})| > \varepsilon\} = 0 \tag{2}$$

(cf. théorème 4, chapitre IV, § 5). De (2) il découle qu'existe une suite $\varepsilon_n \downarrow 0$ telle que

$$\lim_{n\to\infty} \sum_{k=1}^{n} \mathsf{P}\{|\xi(t_{nk}) - \xi(t_{n,k-1})| > \varepsilon_n\} = 0. \tag{3}$$

Posons

$$\xi_{nk} = \begin{cases} \xi(t_{nk}) - \xi(t_{n,k-1}) & \text{si} \quad |\xi(t_{nk}) - \xi(t_{n,k-1})| \leqslant \varepsilon_n, \\ 0 & \text{si} \quad |\xi(t_{nk}) - \xi(t_{n,k-1})| > \varepsilon_n. \end{cases}$$

Alors

$$\mathsf{P}\left\{\xi(t) \neq \sum_{k=1}^{n} \xi_{nk}\right\} \leqslant \sum_{k=1}^{n} \mathsf{P}\{|\xi(t_{nk}) - \xi(t_{n,k-1})| > \varepsilon_n\}$$

et par suite $\sum_{k=1}^{n} \xi_{nk}$ converge stochastiquement vers $\xi(t)$.

Le théorème limite central dit que

$$\left(\sum_{k=1}^{n} \xi_{nk} - \mathsf{E} \sum_{k=1}^{n} \xi_{nk}\right) / \mathrm{Var} \sum_{k=1}^{n} \xi_{nk}$$

admet une répartition normale limite non dégénérée pourvu que $\varliminf_{n\to\infty} \mathrm{Var} \sum_{k=1}^{n} \xi_{nk} > 0$. Alors

$$\left(\xi(t) - \mathsf{E} \sum_{k=1}^{n} \xi_{nk}\right) / \mathrm{Var} \sum_{k=1}^{n} \xi_{nk}$$

admettra une répartition normale limite non dégénérée, ce qui n'est possible que si $\xi(t)$ possède une répartition normale. Si pour une suite n_r

$$\lim_{r\to\infty} \mathrm{Var} \sum_{k=1}^{n_r} \xi_{n_r k} \to 0,$$

alors

$$\sum_{k=1}^{n_r} \xi_{n_r k} - \mathsf{E} \sum_{k=1}^{n_r} \xi_{n_r k} \to 0$$

stochastiquement, et

$$\xi(t) - \mathsf{E}\sum_{k=1}^{n_r} \xi_{n_r k} \to 0$$

stochastiquement, ce qui n'a lieu que si $\xi(t)$ est presque sûrement constant. ■

Si $\xi(t)$ est un processus continu homogène, il existe $a \in \mathcal{R}^m$ et $B \in \mathcal{L}_+(\mathcal{R}^m)$ tels que les fonctions $a(t)$ et $B(t)$ de la formule (1) sont de la forme

$$a(t) = ta, \quad B(t) = tB.$$

Un processus homogène $w(t)$ tel que $a = 0$, $B = I$ (I est l'opérateur unitaire) est dit wienérien. Si $B^{1/2}$ désigne la racine carrée (non négative) de l'opérateur non négatif B, alors le processus

$$\xi(t) = ta + B^{1/2} w(t),$$

où $w(t)$ est un processus wienérien, sera processus gaussien homogène à accroissements indépendants et tel que

$$\mathsf{E}\exp\{i(z, \xi(t))\} = \exp\left\{t(z, a) - \frac{1}{2}t(Bz, z)\right\}$$

Si la fonction $B(t)$ est dérivable, on peut exprimer un processus de fonction caractéristique (1) au moyen d'un processus wienérien. Pour cela nous aurons besoin d'intégrales stochastiques par rapport au processus wienérien de la forme

$$\int_0^T Z(t)\, dw(t), \tag{4}$$

où $Z(t)$ est une fonction opérationnelle (non aléatoire). Choisissons dans $\mathcal{R}^m$ une base orthonormée $\{e_1, \ldots, e_m\}$. Posons $w_k(t) = (e_k, w(t))$. Comme

$$\mathsf{E}\exp\left\{i\sum_{k=1}^m \lambda_k w_k(t)\right\} = \mathsf{E}\exp\left\{i\left(\sum_{k=1}^m \lambda_k e_k, w(t)\right)\right\} =$$
$$= \exp\left\{-\frac{1}{2}\left(\sum_{k=1}^m \lambda_k e_k, \sum_{k=1}^m \lambda_k e_k\right)\right\} = \exp\left\{-\frac{1}{2}\sum_{k=1}^m \lambda_k^2\right\},$$

les processus $w_k(t)$ sont des processus wienériens à une dimension mutuellement indépendants. Chacun d'eux est visiblement à accroissements orthogonaux : pour $t_1 < t_2 < t_3$

$$\mathsf{E}(w_k(t_3) - w_k(t_2))(w_k(t_2) - w_k(t_1)) =$$
$$= \mathsf{E}(w_k(t_3) - w_k(t_2))\,\mathsf{E}(w_k(t_2) - w_k(t_1)) = 0,$$

et de plus

$$\mathsf{E}\,|w_k(t_2) - w_k(t_1)|^2 = t_2 - t_1.$$

Les intégrales

$$\int_0^T f(s)\, dw_k(s)$$

sont définies pour tous les $f \in \mathcal{L}_2\,[0,\ T]$ (cf. chapitre V, § 3).

L'intégrale (4) se détermine au moyen de l'expression

$$\int_0^T Z(t)\, dw(t) = \sum_{k=1}^{m} \Big(\sum_{j=1}^{m} \int_0^T (Z(t)\, e_j,\ e_k)\, dw_j(t) \Big)\, e_k. \tag{5}$$

Cette intégrale est définie pour toutes les fonctions opérationnelles mesurables $Z(t)$ telles que

$$\int_0^T \operatorname{Tr} Z(t)\, Z^*(t)\, dt < \infty,$$

Z^* est le conjugué de Z, Tr ZZ^* la trace de l'opérateur ZZ^*.

L'indépendance des processus $w_j(t)$ et (5) montrent que

$$\left.\begin{aligned} &\mathsf{E} \int_0^T Z(t)\, dw(t) = 0, \\ &\mathsf{E} \Big(\int_0^T Z_1(t)\, dw(t),\ \int_0^T Z_2(t)\, dw(t) \Big) = \int_0^T \operatorname{Tr} Z_1(t)\, Z_2^*(t)\, dt, \end{aligned}\right\} \tag{6}$$

$$\mathsf{E} \Big(\int_0^T Z_1(t)\, dw(t),\ z \Big)^2 = \int_0^T (Z_1(t)\, Z_1^*(t)\, z,\ z)\, dt. \tag{7}$$

A noter que l'intégrale (5) possède une répartition gaussienne (puisque elle est limite d'intégrales de fonctions simples s'exprimant linéairement en fonction des accroissements d'un processus wienérien et admettant donc une répartition gaussienne).

Soit maintenant $C(t) = \Big(\frac{d}{dt} B(t)\Big)^{1/2}$. Le processus

$$\xi(t) = a(t) + \int_0^t C(s)\, dw(s)$$

est gaussien; en vertu de (6) et (7) on a

$$\mathsf{E}\, \xi(t) = a(t),$$

$$\mathsf{E}\, (\xi(t) - a(t),\ z)^2 = \mathsf{E} \Big(\int_0^t C(s)\, dw(s),\ z \Big)^2 =$$

$$= \int_0^t \left(\left(\frac{d}{ds} B(s) \right)^{1/2} \left(\frac{d}{ds} B(s) \right)^{1/2} z, \ z \right) ds =$$

$$= \int_0^t \left(\frac{d}{ds} B(s) z, \ z \right) = (B(t) z, \ z) - (B(0) z, \ z).$$

Donc l'accroissement du processus $\xi(t)$ possède la fonction caractéristique (1).

A noter qu'il est toujours possible d'exhiber une fonction $\lambda(t)$ strictement croissante telle que $B(t)$ soit absolument continue en $\lambda(t)$. Pour une telle fonction on peut par exemple prendre $\lambda(t) = t + \mathrm{Tr}\,[B(t) - B(0)]$. Comme l'opérateur non négatif B est tel que $\| B \| \leqslant \mathrm{Tr}\, B$, pour $t_1 < t_2$ on a

$$\| B(t_2) - B(t_1) \| \leqslant \mathrm{Tr}\,[B(t_2) - B(t_1)] \leqslant \lambda(t_2) - \lambda(t_1). \quad (8)$$

Si à présent la fonction $\varphi(t)$ est l'inverse de $\lambda(t)$: $\lambda(\varphi(t)) = t$, alors la fonction caractéristique de l'accroissement du processus $\hat{\xi}(t) = \xi(\varphi(t))$ est:

$$\mathsf{E} \exp \{ i (z, \ \hat{\xi}(t_2) - \hat{\xi}(t_1)) \} =$$

$$= \exp \left\{ i (z, \ \hat{a}(t_2) - \hat{a}(t_1)) - \frac{1}{2} ((\hat{B}(t_2) - \hat{B}(t_1)) z, \ z) \right\},$$

où $\hat{a}(t) = a(\varphi(t))$, $\hat{B}(t) = B(\varphi(t))$, et de plus en vertu de (8)

$$\| \hat{B}(t_2) - \hat{B}(t_1) \| \leqslant \lambda(\varphi(t_2) - \lambda(\varphi(t_1)) = t_2 - t_1,$$

i.e. $\hat{B}(t)$ est absolument continu. Comme $\xi(t) = \hat{\xi}(\lambda(t))$, on peut affirmer que tout processus continu à accroissements indépendants est susceptible d'être déduit à partir de la somme d'une fonction continue non aléatoire et d'une intégrale stochastique de la forme (5) par un changement de temps continu monotone (non aléatoire).

Le processus wienérien s'appelle également *processus du mouvement brownien*. Ceci pour la raison suivante. Etudions le mouvement d'une particule assez petite dans un liquide, sous l'influence des chocs avec les molécules du liquide en agitation thermique. En physique ce phénomène porte le nom de « mouvement brownien ».

On conviendra dans l'étude probabiliste de ce phénomène que les vitesses des molécules heurtées par la particule sont aléatoires et que dans un liquide homogène la répartition de la vitesse ne dépend pas de la position de la molécule (mais seulement de la température qui, elle, est partout la même). Si l'on admet ensuite que les vitesses des molécules sont indépendantes entre elles et que l'on néglige l'inertie de la particule, le déplacement de la particule d'un point à un autre, pendant un certain intervalle de temps, ne dépendra pas de la position de la particule et de son mouvement passé. Donc le

processus $\xi(t)$ qui décrit la position de la particule dans $\mathcal{R}^3$ à l'instant t sera un processus à accroissements indépendants. D'autre part, des considérations d'ordre physique font ressortir manifestement que ce processus sera continu et homogène dans le temps si l'état physique du liquide ne varie pas. Or tout processus homogène continu à accroissements indépendants est gaussien. On supposera qu'à l'instant initial la particule se trouve à l'origine des coordonnées, i.e. $\xi(0) = 0$. Soient $\mathsf{E}\xi(t) = a(t)$, $\mathrm{Var}(\xi(t), z) = (Bz, z)$, $a(t)$ vecteur de $\mathcal{R}^3$, B opérateur symétrique dans $\mathcal{R}^3$. Dans le cas d'un liquide homogène au repos, le processus doit être isotrope (car la répartition des projections de la vitesse des molécules du liquide sur une direction quelconque ne dépend pas de cette direction), i.e. (a, z) et (Bz, z) ne doivent pas dépendre de z pour $|z| = 1$. Ceci n'est possible que si $(a, z) = 0$, $(Bz, z) = c(z, z)$. Ainsi les considérations les plus générales nous ont conduits au processus du mouvement brownien défini plus haut.

Etudions en détail le processus wienérien à une dimension.

Soit $a \neq 0$ un nombre. Appelons τ_a un instant tel que $\frac{w(t)}{a} \leqslant 1$ pour $t \leqslant \tau_a$ et $\sup\limits_{\tau_a \leqslant t \leqslant \tau_a + \delta} \frac{w(t)}{a} > 1$ $\forall \delta > 0$. Si $\frac{w(t)}{a} \leqslant 1$ pour tous les t, on admettra que $\tau_a = +\infty$. L'instant τ_a sera dit instant de première intersection du niveau a par le processus $w(t)$.

Soit maintenant τ'_a un instant tel que $\frac{w(t)}{a} < 1$ pour $t < \tau'_a$, $w(\tau'_a) = a$. L'instant τ'_a sera appelé instant de premier accès du processus $w(t)$ au niveau a. De toute évidence $\tau'_a \leqslant \tau_a$. On a le

LEMME 1. $\mathsf{P}\{\tau'_a = \tau_a\} = 1$.

Démonstration. Le processus $w(t)$ étant symétrique (car $-w(t)$ a les mêmes répartitions), on admet que $a > 0$. L'événement $\{\tau'_a < \tau_a\}$ implique l'un au moins des événements

$$\{\max_{0 \leqslant s \leqslant \frac{r}{m}} w(s) = a\}, \quad r = 1, 2, \ldots, \quad m = 1, 2, \ldots, \quad r < mt.$$

Donc la démonstration du lemme passe par celle de

$$\mathsf{P}\{\max_{0 \leqslant s \leqslant t} w(s) = a\} = 0$$

pour $a > 0$ et $\forall t$; il est immédiat que pour $t_1 < t$

$$\mathsf{P}\{\max_{0 \leqslant s \leqslant t} w(s) = a\} \leqslant \mathsf{P}\{\max_{0 \leqslant s \leqslant t_1} w(s) = a\} + \\ + \int_{-\infty}^{a} \mathsf{P}\{w(t_1) \in dx\}\, \mathsf{P}\{\max_{t_1 \leqslant s \leqslant t} w(s) - w(t_1) = a - x\}.$$

Or $\mathsf{P}\{\max_{t_1\leqslant s\leqslant t} w(s)-w(t_1)=z\}$ est non nul pour un nombre au plus dénombrable de z, i.e. $\mathsf{P}\{\max_{t_1\leqslant s\leqslant t} w(s)-w(t_1)=a-x\}=0$ pour presque tous les x (par rapport à la mesure de Lebesgue). Par suite

$$\int_{-\infty}^{a} \mathsf{P}\{w(t_1)\in dx\}\,\mathsf{P}\{\max_{t_1\leqslant s\leqslant t} w(s)-w(t_1)=a-x\}=$$

$$=\frac{1}{\sqrt{2\pi t_1}}\int_{-\infty}^{a} \mathsf{P}\{\max_{t_1\leqslant s\leqslant t} w(s)-w(t_1)=a-x\}\,e^{-\frac{x^2}{2t_1}}\,dx=0,$$

puisque la fonction à intégrer est nulle presque partout. Donc

$$\mathsf{P}\{\max_{0\leqslant s\leqslant t} w(s)=a\}\leqslant \mathsf{P}\{\max_{0\leqslant s\leqslant t_1} w(s)=a\},$$

i.e. $\mathsf{P}\{\max_{0\leqslant s\leqslant t} w(s)=a\}$ ne décroît pas pour $t\downarrow 0$ et en même temps la continuité de $w(t)$ entraîne que $\mathsf{P}\{\max_{0\leqslant s\leqslant t} w(s)>\varepsilon\}\to 0$ avec t pour $\varepsilon>0$. Donc

$$\mathsf{P}\{\max_{0\leqslant s\leqslant t} w(s)=a\}\leqslant \lim_{t_1\downarrow 0} \mathsf{P}\left\{\max_{0\leqslant s\leqslant t} w(s)>\frac{a}{2}\right\}=0. \blacksquare$$

Ce lemme nous permettra dans la suite de ne pas faire de distinction entre l'instant de premier accès et l'instant de première intersection du niveau a. Nous noterons τ_a ces deux instants.

Pour étudier certaines caractéristiques du processus $w(t)$ nous aurons besoin du lemme suivant.

LEMME 2. *Soient $w(t)$ un processus du mouvement brownien, $a\neq 0$, τ_a l'instant de première intersection du niveau a par $w(t)$, $w_1(t)$ un processus tel que $w_1(t)=w(t)$ pour $t<\tau_a$ et $w_1(t)=2a-w(t)$ pour $t\geqslant\tau_a$. Le processus $w_1(t)$ est également wienérien.*

Démonstration. Posons

$$w_{nk}=w\left(\frac{k}{n}\right)-w\left(\frac{k-1}{n}\right),$$

$$w^{(n)}(t)=\sum_{k\leqslant nt} w_{nk},$$

$$w_1^{(n)}(t)=\sum_{k\leqslant nt} (-1)^{\varepsilon_{nk}}\, w_{nk},$$

où ε_{nk} est égal à 0 ou 1 selon que $\sup_{j\leqslant k-1} \dfrac{w^{(n)}\left(\frac{j}{n}\right)}{a}$ est $\leqslant 1$ ou >1. On remarquera que les quantités $(-1)^{\varepsilon_{nk}} w_{nk}$ sont indépendantes et équiréparties pour $k=1, 2, \ldots$ et que leurs répartitions sont

confondues avec celles des w_{nk}, car w_{nk} et $-w_{nk}$ sont équiréparties et les w_{nk} ne dépendent pas de ε_{nk}, $w_{n1}, \ldots, w_{n,\, k-1}$. Donc $w_{nk}(-1)^{\varepsilon_{nk}}$ admettent une répartition normale d'espérance mathématique nulle et de variance $1/n$. Par suite, les répartitions finidimensionnelles des processus $w^{(n)}(t)$ et $w_1^{(n)}(t)$ sont confondues. La démonstration de ce lemme résulte de ce que $w^{(n)}(t) \to w(t)$, $w_1^{(n)}(t) \to w_1(t)$ presque sûrement, puisque le processus du mouvement brownien est continu.

Appliquons le lemme démontré à la recherche des caractéristiques suivantes du processus: $\max\limits_{0 \leqslant t \leqslant T} w(t)$, $\min\limits_{0 \leqslant t \leqslant T} w(t)$, $\max\limits_{0 \leqslant t \leqslant T} |w(t)|$.

THÉORÈME 2. *Pour* $a > 0$

$$\mathsf{P}\{\max_{0 \leqslant t \leqslant T} w(t) > a,\ w(T) \in [c, d]\} =$$

$$= \frac{1}{\sqrt{2\pi T}} \int_{c \vee a}^{d \vee a} e^{-\frac{x^2}{2T}} dx + \frac{1}{\sqrt{2\pi T}} \int_{(2a-d) \vee a}^{(2a-c) \vee a} e^{-\frac{x^2}{2T}} dx, \qquad (9)$$

où $a \vee b = \max[a, b]$.

Démonstration. Utilisons la relation

$$\mathsf{P}\{\max_{0 \leqslant t \leqslant T} w(t) > a,\ w(T) \in [c, d]\} = \mathsf{P}\{w(T) \in [c, d] \cap]a, \infty[\} +$$
$$+ \mathsf{P}\{\max_{0 \leqslant t \leqslant T} w(t) > a,\ w(T) \in [c, d] \cap]-\infty, a]\}$$

(qui est vraie puisque l'événement $\{w(T) \in [c, d] \cap]a, \infty[\}$ implique l'événement $\{\max\limits_{a \leqslant t \leqslant T} w(t) > a\}$). Maintenant nous allons chercher la probabilité

$$\mathsf{P}\{\max_{0 \leqslant t \leqslant T} w(t) > a,\ w(T) \in [c, d] \cap]-\infty, a]\}.$$

Soit $w_1(t)$ un processus défini comme celui du lemme 2. L'événement $\{\max\limits_{0 \leqslant t \leqslant T} w(t) > a,\ w(T) \in [c, d] \cap]-\infty, a]\}$ coïncide avec l'événement $\{\max\limits_{0 \leqslant t \leqslant T} w_1(t) \geqslant a,\ w_1(T) \in [2a-d, 2a-c] \cap [a, \infty[\}$. Or, l'événement $\{w_1(T) \in [2a-d, 2a-c] \cap [a, \infty[\}$ implique l'événement $\{\max\limits_{0 \leqslant t \leqslant T} w_1(t) \geqslant a\}$, donc

$$\{\max_{0 \leqslant t \leqslant T} w_1(t) \geqslant a,\ w_1(T) \in [2a-d, 2a-c] \cap [a, \infty[\} =$$
$$= \{w_1(T) \in [2a-d, 2a-c] \cap [a, \infty[\}.$$

Par suite, en vertu du lemme 2

$$\mathsf{P}\{\max_{0 \leqslant t \leqslant T} w(t) > a,\ w(T) \in [c, d] \cap]-\infty, a]\} =$$

$$=\frac{1}{\sqrt{2\pi T}}\int_{(2a-d)\vee a}^{(2a-c)\vee a} e^{-\frac{x^2}{2T}}\,dx. \qquad (10)$$

D'autre part,

$$\mathsf{P}\{w(T)\in[c,\ d]\cap[a,\ \infty[\,\}=\int_{c\vee a}^{d\vee a}\frac{1}{\sqrt{2\pi T}}e^{-\frac{x^2}{2T}}\,dx. \qquad (11)$$

(10) et (11) achèvent la démonstration du théorème.

COROLLAIRE. *Pour* $a>0$

$$\mathsf{P}\{\max_{0\leqslant t\leqslant T} w(t)>a\}=\frac{2}{\sqrt{2\pi T}}\int_0^\infty e^{-\frac{x^2}{2T}}\,dx.$$

Ceci découle du théorème (2) si $]c,\ d]=]-\infty,\ \infty[$.

THÉORÈME 3. *Soient* $a_1<0<a_2$ *et* $[c,\ d]\subset[a_1,\ a_2]$. *Alors*

$$\mathsf{P}\{\min_{0\leqslant t\leqslant T} w(t)>a_1,\ \max_{0\leqslant t\leqslant T} w(t)<a_2,\ w(T)\in[c,\ d]\}=$$

$$=\frac{1}{\sqrt{2\pi T}}\sum_{k=-\infty}^{\infty}\int_c^d\left[\exp\left\{-\frac{1}{2T}(x+2k(a_2-a_1))^2\right\}-\right.$$

$$\left.-\exp\left\{-\frac{1}{2T}(x-2a_2+2k(a_2-a_1))^2\right\}\right]dx. \qquad (12)$$

D é m o n s t r a t i o n. Appelons $\mathfrak{A}_k^{(i)}$ l'événement: le processus $w(t)$ coupe le niveau a_i avant le niveau a_j ($j\neq i,\ i,\ j=1,\ 2$) sur l'intervalle $[0,\ T]$ et puis au moins k fois l'intervalle $[a_1,\ a_2]$ (on admet que la fonction $x(t)$ coupe k fois l'intervalle $[a_1,\ a_2]$ si la fonction $\operatorname{sgn}(x(t)-a_1)+\operatorname{sgn}(x(t)-a_2)$ change k fois de signe) et $w(T)\in[c,\ d]$. La probabilité cherchée se note:

$$\mathsf{P}\{w(T)\in[c,\ d]\}-\mathsf{P}\{\mathfrak{A}_0^{(1)}\}-\mathsf{P}\{\mathfrak{A}_0^{(2)}\}.$$

Pour calculer $\mathsf{P}\{\mathfrak{A}_0^{(i)}\}$ il faut connaître

$$\mathsf{P}\{\mathfrak{A}_k^{(i)}\}+\mathsf{P}\{\mathfrak{A}_{k+1}^{(j)}\}=\mathsf{P}\{\mathfrak{A}_k^{(i)}\cup\mathfrak{A}_{k+1}^{(j)}\}\quad(i\neq j,\ i,\ j=1,\ 2).$$

Il est aisé de voir que l'événement $\mathfrak{A}_k^{(i)}\cup\mathfrak{A}_{k+1}^{(j)}$ traduit le fait qu'avant l'instant T le processus $w(t)$ coupe le niveau a_i (pas forcément avant de couper a_j), ensuite l'intervalle $[a_1,\ a_2]$ au moins k fois et tombe dans l'intervalle $[c,\ d]$ pour $t=T$. Soient τ_1 l'instant de première intersection du niveau a_i, τ_2 l'instant de première intersection de a_j après l'instant τ_1, τ_3 l'instant de première intersection de a_i après τ_2, etc. Posons

$$w_1(t)=\begin{cases} w(t) & \text{pour } t<\tau_1,\\ 2w(\tau_1)-w(t) & \text{pour } t\geqslant\tau_1,\end{cases}$$

$$w_2(t) = \begin{cases} w_1(t) & \text{pour } t < \tau_2, \\ 2w(\tau_2) - w_1(t) & \text{pour } t \geqslant \tau_2, \end{cases}$$

$$w_3(t) = \begin{cases} w_2(t) & \text{pour } t < \tau_3, \\ 2w_2(\tau_3) - w_2(t) & \text{pour } t \geqslant \tau_3, \text{ etc.} \end{cases}$$

On remarquera que les processus $w_l(t)$ seront de mouvement brownien, car τ_l est l'instant de première intersection du niveau

$$a_i + (l-1)(a_i - a_j)$$

par le processus $w_{l-1}(t)$. Si l'événement $\mathfrak{A}_k^{(i)} \cup \mathfrak{A}_{k+1}^{(j)}$ a lieu, le processus $w_{k+1}(t)$ $(t < T)$ coupe successivement les niveaux

$$a_i,\ a_i + (a_i - a_j),\ \ldots,\ a_i + k(a_i - a_j)$$

et à l'instant T tombe dans l'intervalle $[c_k, d_k]$, où

$$\left.\begin{aligned} c_k &= c + (k+1)(a_i - a_j), \\ d_k &= d + (k+1)(a_i - a_j) \end{aligned}\right\} \text{ pour } k \text{ impair};$$

$$\left.\begin{aligned} c_k &= 2a_i - d + k(a_i - a_j), \\ d_k &= 2a_i - c + k(a_i - a_j) \end{aligned}\right\} \text{ pour } k \text{ pair}.$$

Si d'autre part $w_{k+1}(t)$ vérifie les conditions énumérées, l'événement $\mathfrak{A}_k^{(i)} \cup \mathfrak{A}_{k+1}^{(j)}$ a lieu. Comme $w_{k+1}(t)$ est un processus continu nul pour $t = 0$, pour qu'il tombe dans l'intervalle $[c_k, d_k]$ il suffit qu'il coupe successivement les niveaux $a_i + l(a_i - a_j)$, $l = 0, \ldots, k$. Donc

$$\mathsf{P}\{\mathfrak{A}_k^{(i)} \cup \mathfrak{A}_{k+1}^{(j)}\} = \mathsf{P}\{w_{k+1}(T) \in [c_k, d_k]\} = \mathsf{P}\{w(T) \in [c_k, d_k]\}.$$

Le processus $w(t)$ étant continu, il coupera presque sûrement l'intervalle $[a_1, a_2]$ un nombre fini de fois et par suite $\mathsf{P}\{\mathfrak{A}_k^{(i)}\} \to 0$ pour $k \to \infty$. En passant à la limite pour $n \to \infty$ dans

$$\mathsf{P}\{\mathfrak{A}_0^{(1)}\} + \mathsf{P}\{\mathfrak{A}_0^{(2)}\} = (-1)^{n+1}[\mathsf{P}\{\mathfrak{A}_{n+1}^{(1)}\} + \mathsf{P}\{\mathfrak{A}_{n+1}^{(2)}\}] + \\ + \sum_{k=0}^{n} (-1)^k (\mathsf{P}\{\mathfrak{A}_k^{(1)}\} + \mathsf{P}\{\mathfrak{A}_k^{(2)}\} + \mathsf{P}\{\mathfrak{A}_{k+1}^{(1)}\} + \mathsf{P}\{\mathfrak{A}_{k+1}^{(2)}\}),$$

on trouve

$$\mathsf{P}\{\mathfrak{A}_0^{(1)}\} + \mathsf{P}\{\mathfrak{A}_0^{(2)}\} = \sum_{k=0}^{\infty} (-1)^k \sum_{i=1}^{2} (\mathsf{P}\{\mathfrak{A}_k^{(i)}\} + \mathsf{P}\{\mathfrak{A}_{k+1}^{(i)}\}) =$$

$$= \frac{1}{\sqrt{2\pi T}} \sum_{k=0}^{\infty} \Bigg[\int_{2a_1 - d + 2k(a_2 - a_1)}^{2a_1 - c + 2k(a_2 - a_1)} \exp\left(-\frac{x^2}{2T}\right) dx +$$

$$+ \int_{2a_2 - d + 2k(a_2 - a_1)}^{2a_2 - c + 2k(a_2 - a_1)} \exp\left\{-\frac{x^2}{2T}\right\} dx - \int_{c - 2(k+1)(a_2 - a_1)}^{d - 2(k+1)(a_2 - a_1)} \exp\left\{-\frac{x^2}{2T}\right\} dx -$$

$$- \int\limits_{c+2(k+1)(a_2-a_1)}^{d+2(k+1)(a_2-a_1)} \exp\left\{-\frac{x^2}{2T}\right\} dx \Bigg].$$

La probabilité cherchée vaut donc

$$\frac{1}{\sqrt{2\pi T}} \sum_{k=-\infty}^{\infty} \Bigg[\int\limits_{c+2k(a_2-a_1)}^{d+2k(a_2-a_1)} \exp\left\{-\frac{x^2}{2T}\right\} dx - $$

$$- \int\limits_{2a_2-d+2k(a_2-a_1)}^{2a_2-c+2k(a_2-a_1)} \exp\left\{-\frac{x^2}{2T}\right\} dx \Bigg].$$

La substitution $x - 2k(a_2 - a_1) = u$ dans la première intégrale et $2k(a_2 - a_1) + 2a_2 - x = u$ dans la seconde donnent la formule (12). ■

COROLLAIRE 1. *La répartition conjointe de* $\max_{0 \leqslant t \leqslant T} w(t)$ *et de* $\min_{0 \leqslant t \leqslant T} w(t)$ *pour* $a_1 < 0$, $a_2 > 0$ *est définie par*

$$\mathsf{P}\{\min_{0 \leqslant t \leqslant T} w(t) > a_1, \max_{0 \leqslant t \leqslant T} w(t) < a_2\} =$$

$$= \frac{1}{\sqrt{2\pi T}} \sum_{k=-\infty}^{\infty} \int\limits_{a_1}^{a_2} \Big[\exp\left\{-\frac{1}{2T}(x + 2k(a_2 - a_1))^2\right\} -$$

$$- \exp\left\{-\frac{1}{2T}(x - 2a_2 + 2k(a_2 - a_1))^2\right\}\Big] dx. \quad (13)$$

COROLLAIRE 2. *Pour* $a > 0$, $[c, d] \subset [-a, a]$, *on a*

$$\mathsf{P}\{\max_{0 \leqslant t \leqslant T} |w(t)| < a, w(T) \in [c, d]\} =$$

$$= \frac{1}{\sqrt{2\pi T}} \int\limits_c^d \sum_{k=-\infty}^{\infty} (-1)^k \exp\left\{-\frac{1}{2T}(x - 2ka)^2\right\} dx. \quad (14)$$

§ 4. Construction de processus généraux à accroissements indépendants

Soit $\xi(t)$ processus stochastiquement continu séparable à accroissements indépendants, défini pour $t \in [0, T]$ et à valeurs dans X, espace euclidien finidimensionnel. Il présente donc presque sûrement des discontinuités de seconde espèce en vertu du § 4, chapitre IV.

Pour tout $\varepsilon > 0$ il n'existe presque sûrement qu'un nombre fini de points t tels que $|\xi(t + 0) - \xi(t - 0)| > \varepsilon$.

Soient $X_\varepsilon = \{x : |x| > \varepsilon\}$, $\mathfrak{B}_\varepsilon$ la tribu des boréliens de X_ε. De ce qui précède il suit que pour tout $A \in \mathfrak{B}_\varepsilon$ le nombre de points $t \in [0, T]$ tels que $\xi(t + 0) - \xi(t - 0) \in A$ est fini presque sûrement. Appelons $\nu(t, A)$ le nombre de points $s \in [0, t[$ tels que

$\xi(s+0)-\xi(s-0)\in A$. Le processus $\nu(t, A)$ est à accroissements indépendants, puisque $\nu(t_2, A)-\nu(t_1, A)$ $(t_1<t_2)$ est entièrement défini par les accroissements $\xi(s)-\xi(t_1)$ pour $s\in$ $\in[t_1, t_2]$, et par suite les accroissements de $\nu(t, A)$ sur des intervalles disjoints s'expriment en fonction des accroissements de $\xi(t)$ sur des intervalles disjoints. D'autre part, $\nu(t, A)$ sera un processus stochastiquement continu (si $\nu(t', A)-\nu(t, A)$ ne tendait pas presque sûrement vers zéro pour $t'-t\to 0$, la probabilité $\mathsf{P}\{|\xi(t')-\xi(t)|>\varepsilon\}\not\to 0$, ce qui est impossible en vertu de la continuité stochastique de $\xi(t)$).

Comme $\nu(t, A)$ est un processus discontinu stochastiquement continu, tous les sauts sont égaux à l'unité (et partant $\nu(t, A)$ est le nombre de sauts sur l'intervalle $[0, t]$) et $\nu(t, A)$ admet une répartition poissonienne d'après le corollaire du théorème 1, § 2.

Posons $\Pi(t, A)=\mathsf{E}\nu(t, A)$. La fonction d'ensembles $\Pi(t, A)$ (à t fixe) est une mesure sur $\mathfrak{B}_\varepsilon$.

En effet, si $A=\bigcup\limits_{k=1}^{\infty}A_k$ et A_k sont deux à deux disjoints,

$\nu(t, A)=\sum\limits_{k=1}^{\infty}\nu(t, A_k)$ et

$$\mathsf{E}\nu(t, A)=\sum_{k=1}^{\infty}\mathsf{E}\nu(t, A_k)$$

puisque

$$0\leqslant\sum_{k=1}^{n}\nu(t, A_k)\leqslant\nu(t, A).$$

L'étude des propriétés de $\nu(t, A)$ passe par celle du processus $\xi(t, A)$ défini par

$$\xi(t, A)=\sum_{s<t}[\xi(s+0)-\xi(s-0)]\,\chi_A(\xi(s+0)-\xi(s-0)),$$

où $\chi_A(x)$ est l'indicateur de l'ensemble A. Autrement dit, $\xi(t, A)$ est la somme des sauts du processus $\xi(t)$ qui ont eu lieu avant l'instant t et sont tombés dans l'ensemble A. Si $A\in\mathfrak{B}_\varepsilon$, ces sauts sont presque sûrement en nombre fini si bien que $\xi(t, A)$ a un sens. La continuité stochastique de $\nu(t, A)$ implique celle de $\xi(t, A)$. De plus, $\xi(t, A)$ est un processus à accroissements indépendants. L'indépendance des processus $\xi(t, A_1), \ldots, \xi(t, A_k)$, où A_1, $A_2, \ldots, A_k$ sont des événements deux à deux disjoints, qui est une propriété importante, découle du

THÉORÈME 1. *Les processus $\xi(t, A)$ et $\xi(t)-\xi(t, A)$, $A\in$ $\in\mathfrak{B}_\varepsilon$, sont des processus indépendants à accroissements indépendants.*

Démonstration. La continuité stochastique du processus $\xi(t) - \xi(t, A)$ résulte de celle des processus $\xi(t)$ et $\xi(t, A)$. Comme pour le processus $\nu(t, A)$, on peut affirmer que le processus $[\xi(t, A); \xi(t) - \xi(t, A)] \in X \times X$ est également processus à accroissements indépendants. Pour démontrer l'indépendance des processus $\xi(t, A)$ et $\xi(t) - \xi(t, A)$ il suffit donc d'établir que pour $z_1, z_2 \in X$, $s < t$, on a

$$\mathsf{E} \exp\{i(z_1, \xi(t, A) - \xi(s, A)) + i(z_2, \xi(t) - \xi(t, A) - \xi(s) + \xi(s, A))\} = \mathsf{E} \exp\{i(z_1, \xi(t, A) - \xi(s, A))\} \times \mathsf{E} \exp\{i(z_2, \xi(t) - \xi(t, A) - \xi(s) + \xi(s, A))\}. \quad (1)$$

En effet, l'indépendance des accroissements du processus $[\xi(t, A); \xi(t) - \xi(t, A)]$ entraîne

$$\mathsf{E} \exp\left\{i \sum_{k=1}^{n} [(z_1^{(k)}, \xi(t_k, A) - \xi(t_{k-1}, A)) + (z_2^{(k)}, \xi(t_k) - \xi(t_k, A) - \xi(t_{k-1}) + \xi(t_{k-1}, A))]\right\} = \mathsf{E} \exp\left\{i \sum_{k=1}^{n} (z_1^{(k)}, \xi(t_k, A) - \xi(t_{k-1}, A))\right\} \times \mathsf{E} \exp\left\{i \sum_{k=1}^{n} (z_2^{(k)}, \xi(t_k) - \xi(t_k, A) - \xi(t_{k-1}) + \xi(t_{k-1}, A))\right\}$$

quels que soient $0 < t_0 \ldots < t_n = T$, $z_1^{(j)}, z_2^{(k)} \in X$, or cela signifie que les processus $\xi(t, A)$ et $\xi(t) - \xi(t, A)$ sont indépendants. Prouvons la relation (1) d'abord pour le cas où $\Pi(T, \Gamma_A) = 0$, Γ_A étant la frontière de l'ensemble A. A remarquer que dans ce cas le processus $\xi(t)$ ne présente pas presque sûrement de sauts sur Γ_A. Soit $s = t_{n0} < t_{n1} < \ldots < t_{nn} = t$ et

$$\lim_{n \to \infty} \max_k (t_{n,k+1} - t_{nk}) = 0.$$

Si $\chi_A(x)$ est l'indicateur de l'ensemble A, on a presque sûrement

$$\xi(t, A) - \xi(s, A) = \lim_{n \to \infty} \sum_{k=1}^{n} \chi_A(\xi(t_{nk}) - \xi(t_{n,k-1}))[\xi(t_{nk}) - \xi(t_{n,k-1})]. \quad (2)$$

Posons

$$\xi_{nk} = \chi_A(\xi(t_{nk}) - \xi(t_{n,k-1}))[\xi(t_{nk}) - \xi(t_{n,k-1})],$$
$$\eta_{nk} = \xi(t_{nk}) - \xi(t_{n,k-1}) - \xi_{nk}.$$

Pour prouver (1) il suffit, eu égard à (2), de montrer que

$$\lim_{n\to\infty} \left| \mathsf{E}\exp\left\{i\sum_{k=1}^{n}(z_1, \xi_{nk})+i\sum_{k=1}^{n}(z_2, \eta_{nk})\right\} - \right.$$
$$\left. - \mathsf{E}\exp\left\{i\sum_{k=1}^{n}(z_1, \xi_{nk})\right\}\mathsf{E}\exp\left\{i\sum_{k=1}^{n}(z_2, \eta_{nk})\right\}\right| = 0.$$

L'indépendance des couples (ξ_{nk}, η_{nk}) et l'inégalité (pour $|a_k| \leqslant 1$, $|b_k| \leqslant 1$)

$$\left|\prod_{k=1}^{n} a_k - \prod_{k=1}^{n} b_k\right| \leqslant \sum_{k=1}^{n} |a_k - b_k|$$

nous montrent que

$$\left|\mathsf{E}\exp\left\{i\left(\sum_{k=1}^{n}(z_1, \xi_{nk})+\sum_{k=1}^{n}(z_2, \eta_{nk})\right)\right\} - \right.$$
$$\left. - \mathsf{E}\exp\left\{i\sum_{k=1}^{n}(z_1, \xi_{nk})\right\}\mathsf{E}\exp\left\{i\sum_{k=1}^{n}(z_2, \eta_{nk})\right\}\right| =$$
$$= \left|\prod_{k=1}^{n}\mathsf{E}\exp\{i(z_1, \xi_{nk})+i(z_2, \eta_{nk})\} - \right.$$
$$\left. - \prod_{k=1}^{n}\mathsf{E}\exp\{i(z_1, \xi_{nk})\}\mathsf{E}\exp\{i(z_2, \eta_{nk})\}\right| \leqslant$$
$$\leqslant \sum_{k=1}^{n}\left|\mathsf{E}\exp\{i(z_1, \xi_{nk})+i(z_2, \eta_{nk})\} - \right.$$
$$\left. - \mathsf{E}\exp\{i(z_1, \xi_{nk})\}\mathsf{E}\exp\{i(z_2, \eta_{nk})\}\right|.$$

Comme $(z_1, \xi_{nk})(z_2, \eta_{nk}) = 0$ (i.e. ces deux quantités ne peuvent être simultanément distinctes de zéro), il vient

$$e^{i(z_1, \xi_{nk})+i(z_2, \eta_{nk})} = 1 + \sum_{m=1}^{\infty}\frac{(i(z_1, \xi_{nk})+i(z_2, \eta_{nk}))^m}{m!} =$$
$$= 1 + \sum_{m=1}^{\infty}\left[\frac{(i(z_1, \xi_{nk}))^m}{m!}+\frac{(i(z_2, \eta_{nk}))^m}{m!}\right] = e^{i(z_1, \xi_{nk})}+e^{i(z_2, \eta_{nk})}-1.$$

Donc

$$|\mathsf{E}\exp\{i(z_1, \xi_{nk})+i(z_2, \eta_{nk})\} - \mathsf{E}\exp\{i(z_1, \xi_{nk})\}\mathsf{E}\exp\{i(z_2, \eta_{nk})\}| =$$
$$= |\mathsf{E}\exp\{i(z_1, \xi_{nk})\}-1|\,|\mathsf{E}\exp\{i(z_2, \eta_{nk})\}-1|.$$

Or

$$|\mathsf{E}e^{i(z_1, \xi_{nk})}-1| \leqslant \mathsf{E}\,|e^{i(z_1, \xi_{nk})}-1| \leqslant 2\mathsf{P}\{|\xi_{nk}|>0\} =$$
$$= 2\mathsf{P}\{\chi_A(\xi(t_{nk})-\xi(t_{n,k-1}))>0\}$$

et $\forall \rho > 0$

$$|\mathsf{E} e^{i(z_2, \eta_{nk})} - 1| \leqslant \sup_{|x| \leqslant \rho} |1 - e^{i(z_2, x)}| + 2\mathsf{P}\{|\eta_{nk}| > \rho\},$$

et par suite

$$\overline{\lim_{n\to\infty}} \left| \mathsf{E} \exp\left\{i\left(z_1, \sum_{k=1}^n \xi_{nk}\right) + i\left(z_2, \sum_{k=1}^n \eta_{nk}\right)\right\} - \right.$$

$$\left. - \mathsf{E}\exp\left\{i\left(z_1, \sum_{k=1}^n \xi_{nk}\right)\right\} \mathsf{E}\exp\left\{i\left(z_2, \sum_{k=1}^n \eta_{nk}\right)\right\} \right| \leqslant$$

$$\leqslant \overline{\lim_{n\to\infty}} \left(\sup_{|x|\leqslant\rho} |e^{i(z_2, x)} - 1| + 2 \sup_k \mathsf{P}\{|\eta_{nk}| > \rho\}\right) \times$$

$$\times \overline{\lim_{n\to\infty}}\, 2 \sum_{k=1}^n \mathsf{P}\{\chi_A(\xi(t_{nk}) - \xi(t_{n,k-1})) > 0\} =$$

$$= 2\, \overline{\lim_{n\to\infty}} \left(\sup_{|x|\leqslant\rho} |e^{i(z_2, x)} - 1| + 2\sup_k \mathsf{P}\{|\eta_{nk}| > \rho\}\right) \times$$

$$\times \overline{\lim_{n\to\infty}}\, \mathsf{E} \sum_{k=1}^n \chi_A(\xi(t_{nk}) - \xi(t_{n,k-1})).$$

Comme pour $\rho < \varepsilon$ $(A \in \mathfrak{B}_\varepsilon)$ 58

$$\mathsf{P}\{|\eta_{nk}| > \rho\} \leqslant \mathsf{P}\{|\xi(t_{nk}) - \xi(t_{n,k-1})| > \rho\},$$

la continuité stochastique uniforme de $\xi(t)$ entraîne que

$$\overline{\lim_{n\to\infty}}\, \mathsf{P}\{|\eta_{nk}| > \rho\} = 0.$$

Comme

$$\lim_{n\to\infty} \sum_{k=1}^n \chi_A(\xi(t_{nk}) - \xi(t_{n,k-1})) = \nu(t, A) - \nu(s, A)$$

presque sûrement, exactement comme dans la démonstration du théorème 1, § 2, on trouve

$$\overline{\lim_{n\to\infty}}\, \mathsf{E} \sum_{k=1}^n \chi_A(\xi(t_{nk}) - \xi(t_{n,k-1})) \leqslant$$

$$\leqslant -\ln \mathsf{P}\{\nu(t, A) - \nu(s, A) = 0\} = C < \infty.$$

Donc

$$\overline{\lim_{n\to\infty}} \left| \mathsf{E} \exp\left\{i\left(z_1, \sum_{k=1}^n \xi_{nk}\right) + i\left(z_2, \sum_{k=1}^n \eta_{nk}\right)\right\} - \right.$$

$$\left. - \mathsf{E}\exp\left\{i\left(z_1, \sum_{k=1}^n \xi_{nk}\right)\right\} \mathsf{E}\exp\left\{i\left(z_2, \sum_{k=1}^n \eta_{nk}\right)\right\} \right| \leqslant$$

$$\leqslant 2C \sup_{|x|\leqslant\rho} |e^{i(z_2, x)} - 1|.$$

En passant à la limite pour $\rho \to 0$, on obtient (1), ce qui démontre le théorème pour le cas $\Pi(T, \Gamma_A) = 0$.

Avant de généraliser on remarquera que la famille des ensembles de $\mathfrak{A}$, pour lesquels le théorème est vrai, forme une classe monotone (cf. H a l m o s [1], chapitre 1, § 6), car pour toute suite d'ensembles $A_n \in \mathfrak{B}_\varepsilon$ on a presque sûrement

$$\xi(t, \bigcup_n A_n) = \lim_{n\to\infty} \xi(t, \bigcup_{k=1}^n A_k),$$

$$\xi(t, \bigcap_n A_n) = \lim_{n\to\infty} \xi(t, \bigcap_{k=1}^n A_k)$$

et le passage à la limite ne viole pas l'indépendance des variables aléatoires. On constate aussi que si $\varepsilon > 0$ est tel que $\Pi(T, \Gamma_{X_\varepsilon}) = 0$, les ensembles $A \in \mathfrak{B}_\varepsilon$ tels que $\Pi(T, \Gamma_A) = 0$ forment une algèbre d'ensembles. Or, toute algèbre monotone est une tribu (cf. H a l m o s [1], chapitre 1, § 6), donc $\mathfrak{A}$ est une tribu. On remarquera enfin que $\mathfrak{A}$ contient les boules centrées en chaque point de X de rayons aussi petits que l'on veut (puisque les sphères $S_\rho(x)$ de même centre x et de rayons ρ différents ne possèdent pas de points communs mais $\Pi(T, \Gamma_{S_\rho(x)}) > 0$ pour un nombre au plus dénombrable de valeurs de ρ). Donc $\mathfrak{A}$ contient tous les ensembles ouverts de $\mathfrak{B}_\varepsilon$ et $\mathfrak{A} \supset \mathfrak{B}_\varepsilon$. ∎

COROLLAIRE 1. *Si* $A_1, A_2, \ldots, A_k \in \mathfrak{B}_\varepsilon$ *pour* $\varepsilon > 0$ *et sont deux à deux disjoints, les processus* $\xi(t, A_1), \xi(t, A_2), \ldots, \xi(t, A_k)$ *et* $\xi(t) - \sum_{j=1}^k \xi(t, A_j)$ *sont indépendants.*

En effet, le processus

$$\xi(t) - \sum_{j=1}^k \xi(t, A_j) = \xi(t) - \xi(t, \bigcup_{j=1}^k A_j)$$

ne dépend pas du processus $\xi(t, \bigcup_{j=1}^k A_j)$. Les processus $\xi(t, A_j)$ sont complètement définis par le processus $\xi(t, \bigcup_{j=1}^k A_j)$, donc dans leur ensemble ils ne dépendent pas du processus $\xi(t) - \sum_{j=1}^k \xi(t, A_j)$. De façon analogue l'ensemble des processus $\xi(t, A_j)$, $j \neq i$, $\xi(t) - \sum_{j=1}^k \xi(t, A_j)$ est complètement défini par le processus $\xi(t) - \xi(t, A_i)$ qui ne dépend pas de $\xi(t, A_i)$, puisque cet ensemble

n'en dépend pas non plus. Donc parmi les processus

$$\xi(t, A_j),\ j=1, 2, \ldots, k;\quad \xi(t)-\sum_{j=1}^{k}\xi(t, A_j)$$

aucun ne dépend de l'ensemble des autres. D'où le

COROLLAIRE 2. *Les processus* $\nu(t, A_1), \ldots, \nu(t, A_k)$, *où* $A_1, A_2, \ldots, A_k$ *sont des ensembles de* $\mathfrak{B}_\varepsilon$ *deux à deux disjoints, sont indépendants.*

Ceci résulte de la proposition précédente puisque le processus $\nu(t, A)$ est complètement défini par le processus $\xi(t, A)$.

Soit B^* borélien de $[0, T] \times X_\varepsilon$. Appelons $\nu^*(B^*)$ l'ensemble des t tels que le couple $(t;\ \xi(t+0) - \xi(t-0))$ appartienne à B^*. Il est aisé de s'assurer que ν^* est une mesure aléatoire sur $\mathfrak{B}_\varepsilon^*$, tribu des boréliens de $[0, T] \times X_\varepsilon$. Posons $\Pi^*(B^*) = \mathsf{E}\nu^*(B^*)$; $\Pi^*(B^*)$ est une mesure finie sur $\mathfrak{B}_\varepsilon^*$. Entre les mesures $\nu(t, A)$ et ν^* on a la relation évidente

$$\nu^*([t_1, t_2] \times A) = \nu(t_2, A) - \nu(t_1, A).$$

De façon analogue

$$\Pi^*([t_1, t_2] \times A) = \Pi(t_2, A) - \Pi(t_1, A).$$

COROLLAIRE 3. *La mesure* ν^* *est une mesure aléatoire poissonienne à valeurs indépendantes; la fonction caractéristique de* $\nu^*(B^*)$ *est définie par*

$$\mathsf{E}e^{i\lambda\nu^*(B^*)} = \exp\{(e^{i\lambda}-1)\,\Pi^*(B^*)\}.$$

En effet, soit $\mathfrak{A}_0$ l'algèbre des ensembles engendrée par les ensembles de la forme $[t_1, t_2] \times A$, $[t_1, t_2] \subset [0, T]$, $A \in \mathfrak{B}_\varepsilon$. Si $A_1^*, \ldots, A_k^*$ sont des ensembles disjoints de $\mathfrak{A}_0$, on peut exhiber des ensembles disjoints $\Delta_i^* = [t_1^{(i)}, t_2^{(i)}] \times A_i$, $i = 1, \ldots, N$, tels que A_j^* soient des sommes de Δ_i^*. Les ensembles Δ_i^* peuvent être choisis tels que pour divers i les intervalles $[t_1^{(i)}, t_2^{(i)}]$ comme les ensembles A_i soit n'ont pas de points communs, soit sont confondus. Dans ce cas l'indépendance des $\nu^*(\Delta_i^*)$ découle de celle des accroissements de $\nu(t, A_i)$ pour divers A_i et de celle de $\nu(t, A_i)$. L'indépendance des $\nu^*(\Delta_i^*)$ entraîne celle des $\nu^*(A_j^*)$. Ces quantités admettent une répartition de Poisson comme sommes de variables poissoniennes indépendantes. Pour achever la démonstration il suffit de remarquer que $\sigma(\mathfrak{A}_0) = \mathfrak{B}_\varepsilon^*$.

Soit $\xi_\varepsilon(t)$ le processus déduit de $\xi(t)$ par élimination des sauts supérieurs à ε en valeur absolue: $\xi_\varepsilon(t) = \xi(t) - \xi(t, X_\varepsilon)$. Le processus $\xi_\varepsilon(t)$ sera un processus stochastiquement continu à accroissements indépendants et de sauts $\leqslant \varepsilon$. Tout laisse à croire que $\xi_\varepsilon(t)$ convergera vers un processus continu à accroissements indépendants pour $\varepsilon \to 0$. En fait, ceci est vrai si l'on soustrait de $\xi_\varepsilon(t)$ des

fonctions continues non aléatoires spécialement choisies. Pour le prouver on aura besoin du

LEMME. *Soit* $\xi_\varepsilon(0) = 0$; *alors* $\mathsf{E}\,|\,\xi_\varepsilon(t)\,|^2 < \infty$.

Démonstration. Soit $0 = t_{n0} < t_{n1} < \ldots < t_{nn} = t$ et $\lim\limits_{n\to\infty} \max\limits_k (t_{nk} - t_{n,\,k-1}) = 0$. Posons

$$\xi_{nk} = \psi_{2\varepsilon}(\xi_\varepsilon(t_{nk}) - \xi_\varepsilon(t_{n,\,k-1})),$$

où

$$\psi_\alpha(x) = \begin{cases} x \text{ pour } |x| \leqslant \alpha, \\ 0 \text{ pour } |x| > \alpha. \end{cases}$$

Il est aisé de voir que presque sûrement

$$\xi_\varepsilon(t) = \lim_{n\to\infty} \sum_{k=1}^{n} \xi_{nk}. \tag{3}$$

On remarquera que tous les termes de cette somme sont $\leqslant 2\varepsilon$ en valeur absolue. Soit $\{x_1, \ldots, x_r\}$ une base orthonormée de X. Si $\sum\limits_{k=1}^{n} \operatorname{Var}(\xi_{nk}, x_i)$ était non bornée pour un i, on pourrait exhiber une suite d'indices n telle que $\sum\limits_{k=1}^{n} \operatorname{Var}(\xi_{nk}, x_i) \to \infty$. Dans ce cas

$$\eta_{nk} = \frac{(\xi_{nk} - \mathsf{E}\xi_{nk}, x_i)}{\sqrt{\sum\limits_{j=1}^{n} \operatorname{Var}(\xi_{nj}, x_i)}}$$

rempliraient les conditions du théorème limite central. La somme $\sum\limits_{k=1}^{n} \eta_{nk}$ admettrait une répartition limite, de sorte que $\forall\alpha$

$$\lim_{n\to\infty} \mathsf{P}\left\{\sum_{k=1}^{n} (\xi_{nk}, x_i) > \alpha \sqrt{\sum_{k=1}^{n} \operatorname{Var}(\xi_{nk}, x_i)} + \sum_{k=1}^{n} (\mathsf{E}\xi_{nk}, x_i)\right\} = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\infty} e^{-u^2/2}\, du,$$

$$\lim_{n\to\infty} \mathsf{P}\left\{\sum_{k=1}^{n} (\xi_{nk}, x_i) < -\alpha \sqrt{\sum_{k=1}^{n} \operatorname{Var}(\xi_{nk}, x_i)} + \sum_{k=1}^{n} (\mathsf{E}\xi_{nk}, x_i)\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\alpha} e^{-u^2/2}\, du.$$

La dernière relation contredit le fait, résultant de (3), que $\sum_{k=1}^{n}(\xi_{nk}, x_i)$ est stochastiquement bornée. Donc $\sum_{k=1}^{n}\operatorname{Var}(\xi_{nk}, x_i)$ est bornée pour tous les i.

L'inégalité de Tchébychev nous dit par ailleurs que

$$\mathsf{P}\left\{\left|\sum_{k=1}^{n}(\xi_{nk}, x_i)-\sum_{k=1}^{n}\mathsf{E}(\xi_{nk}, x_i)\right|>L\right\}\leqslant\frac{\operatorname{Var}\sum_{k=1}^{n}(\xi_{nk}, x_i)}{L^2}.$$

Ceci et le fait que $\sum_{k=1}^{n}(\xi_{nk}, x_i)$ est stochastiquement bornée entraînent que $\mathsf{E}\sum_{k=1}^{n}(\xi_{nk}, x_i)$ est bornée. Comme

$$\mathsf{E}\left|\sum_{k=1}^{n}\xi_{nk}\right|^2=\mathsf{E}\sum_{i=1}^{r}\left(\sum_{k=1}^{n}(\xi_{nk}, x_i)\right)^2=$$
$$=\sum_{i=1}^{r}\left[\operatorname{Var}\sum_{k=1}^{n}(\xi_{nk}, x_i)+\left(\mathsf{E}\sum_{k=1}^{n}(\xi_{nk}, x_i)\right)^2\right],$$

$\mathsf{E}\left|\sum_{k=1}^{n}\xi_{nk}\right|^2$ est bornée et partant $\mathsf{E}|\xi_\varepsilon(t)|^2$ puisque

$$\mathsf{E}|\xi_\varepsilon(t)|^2\leqslant\overline{\lim_{n\to\infty}}\,\mathsf{E}\left|\sum_{k=1}^{n}\xi_{nk}\right|^2. \blacksquare$$

Soit ε_n suite convergente de façon monotone vers zéro. Appelons Δ_k l'ensemble des x tels que $\varepsilon_k<|x|\leqslant\varepsilon_{k-1}$, $k=2, 3, \ldots$, Δ_1 l'ensemble des x tels que $|x|>\varepsilon_1$. On remarquera que

$$\xi_{\varepsilon_1}(t)=\sum_{k=2}^{m}\xi(t, \Delta_k)+\xi_{\varepsilon_m}(t)$$

et les termes du second membre sont indépendants d'après le corollaire 1 du théorème 1. Donc pour tout x

$$\sum_{k=2}^{m}\operatorname{Var}(\xi(t, \Delta_k), x)\leqslant\operatorname{Var}(\xi_{\varepsilon_1}(t), x),$$

et par suite la série $\sum_{k=2}^{\infty}\operatorname{Var}(\xi(t, \Delta_k), x)$, $\forall x$, est convergente. Choisissons une suite n_k $(n_1=2)$ telle que

$$\sum_{j=n_k}^{\infty}\operatorname{Var}(\xi(T, \Delta_j), x_i)\leqslant\frac{1}{k^6}\quad\text{pour } i=1, 2, \ldots, r.$$

La suite

$$\sum_{j=2}^{n_k} [\xi(t, \Delta_j) - \mathsf{E}\xi(t, \Delta_j)] \tag{4}$$

sera alors presque sûrement uniformément convergente pour $k \to \infty$. En effet,

$$\mathsf{P}\left\{\sup_{0 \leqslant t \leqslant T} \left| \sum_{j=2}^{n_{k+1}} [\xi(t, \Delta_j) - \mathsf{E}\xi(t, \Delta_j)] - \right.\right.$$

$$\left.\left. - \sum_{j=2}^{n_k} [\xi(t, \Delta_j) - \mathsf{E}\xi(t, \Delta_j)] \right| > \frac{1}{k^2}\right\} \leqslant$$

$$\leqslant \sum_{i=1}^{r} \mathsf{P}\left\{\sup_{0 \leqslant t \leqslant T} \left| \sum_{j=n_k+1}^{n_{k+1}} (\xi(t, \Delta_j) - \mathsf{E}\xi(t, \Delta_j), x_i) \right| \geqslant \frac{1}{\sqrt{r}\, k^2}\right\} \leqslant$$

$$\leqslant \sum_{i=1}^{r} \varlimsup_{m \to \infty} \mathsf{P}\left\{\sup_{l \leqslant mT} \left| \sum_{j=n_k+1}^{n_{k+1}} \left(\xi\left(\frac{l}{m}, \Delta_j\right) - \mathsf{E}\xi\left(\frac{l}{m}, \Delta_j\right), x_i\right) \right| \geqslant\right.$$

$$\left. \geqslant \frac{1}{\sqrt{r}\, k^2}\right\} \leqslant \sum_{i=1}^{r} \varlimsup_{m \to \infty} r k^4 \mathsf{E} \left| \sum_{j=n_k+1}^{n_{k+1}} (\xi(T, \Delta_j) - \mathsf{E}\xi(T, \Delta_j), x_i) \right|^2 \leqslant \frac{r^2}{k^2}$$

(on a utilisé l'inégalité de Kolmogorov, chapitre III, § 1, et la remarque du théorème 5).

La série $\sum_{k=1}^{\infty} \frac{r^2}{k^2}$ étant convergente, le théorème de Borel-Cantelli nous dit que les termes de la série

$$\sum_{k=1}^{\infty} \sum_{j=n_k+1}^{n_{k+1}} (\xi(t, \Delta_j) - \mathsf{E}\xi(t, \Delta_j))$$

sont majorés presque sûrement à partir d'un certain rang par ceux de la série convergente $\sum_{k=1}^{\infty} \frac{1}{k^2}$. D'où résulte la convergence uniforme presque sûre de la suite (4). Il existe donc un processus $\xi_0(t)$, limite uniforme de la suite

$$\xi_{\varepsilon_1}(t) - \sum_{j=2}^{n_k} [\xi(t, \Delta_j) - \mathsf{E}\xi(t, \Delta_j)].$$

Comme $\xi(t, \Delta_j)$ est un processus stochastiquement continu et $\mathsf{E}\,|\,\xi(t, \Delta_j)\,|^2 \leqslant \mathsf{E}\,|\,\xi(T, \Delta_j)\,|^2 < \infty$, le théorème du passage à la limite sous le signe de l'intégrale nous dit que

$$\lim_{t \to s} \mathsf{E}\xi(t, \Delta_j) = \mathsf{E}\xi(s, \Delta_j)$$

et par suite $\mathsf{E}\xi(t, \Delta_j)$ est continue en t. Donc le processus

$$\xi_{\varepsilon_1}(t) - \sum_{j=2}^{n_k} [\xi(t, \Delta_j) - \mathsf{E}\xi(t, \Delta_j)]$$

ne présente presque sûrement pas de sauts $\geqslant \varepsilon_{n_k}$ en valeur absolue et $\xi_0(t)$ (la limite uniforme de ces processus) est presque sûrement continu. On remarquera que la série

$$\sum_{j=2}^{\infty} [\xi(t, \Delta_j) - \mathsf{E}\xi(t, \Delta_j)]$$

est convergente en vertu du théorème de Kolmogorov (chapitre III, § 2, corollaire 2 du théorème 1), car $\sum_{j=2}^{\infty} \mathrm{Var}(\xi(t, \Delta_j), x)$ est convergente pour tout x. La somme de la série $\sum_{j=2}^{\infty} [\xi(t, \Delta_j) - \mathsf{E}\xi(t, \Delta_j)]$ est confondue (mod P) avec celle de la série $\sum_{k=1}^{\infty} \sum_{j=n_k+1}^{n_{k+1}} [\xi(t, \Delta_j) - \mathsf{E}\xi(t, \Delta_j)]$, $\forall t$. D'où le

THÉORÈME 2. *Pour tout processus stochastiquement continu séparable à accroissements indépendants il existe un processus continu $\xi_0(t)$ tel que*

$$\xi(t) = \xi_0(t) + \xi(t, \Delta_1) + \sum_{j=2}^{\infty} [\xi(t, \Delta_j) - \mathsf{E}\xi(t, \Delta_j)].$$

REMARQUE. *Le processus $\xi_0(t)$ comme limite des processus $\xi_{\varepsilon_1}(t) - \sum_{j=2}^{n} [\xi(t, \Delta_j) - \mathsf{E}\xi(t, \Delta_j)]$ ne dépend d'aucun processus $\xi(t, \Delta_j)$, $j = 1, 2, \ldots, n$. Comme*

$$\xi_0(t) + \sum_{j=2}^{\infty} (\xi(t, \Delta_j) - \mathsf{E}\xi(t, \Delta_j)) = \xi_{\varepsilon_1}(t)$$

et $\mathsf{E}\,|\,\xi_{\varepsilon_1}(t)\,|^2 < \infty$ et les termes du second membre sont indépendants, on a $\mathsf{E}\,|\,\xi_0(t)\,|^2 < \infty$.

Considérons les intégrales stochastiques par rapport à la mesure $\nu(t, A)$. On a vu déjà que $\nu(t, A)$ est une fonction d'ensemble A

sur $\mathfrak{B}_\varepsilon$ dénombrablement additive non négative. Soit $\varphi(x)$ fonction mesurable bornée sur tout compact de l'espace X, nulle pour $|x| < \varepsilon$ ($\varepsilon \geqslant 0$). L'intégrale $\int \varphi(x)\, \nu(t, dx)$ se définit comme d'ordinaire. Ceci résulte de la finitude de la mesure $\nu(t, A)$ sur $\mathfrak{B}_\varepsilon$ et aussi du fait que $\nu(t, X_\rho) = 0$, où X_ρ est l'ensemble des x tels que $|x| > \rho$, et

$$\rho = \max_{0 \leqslant s \leqslant t} |\xi(s+0) - \xi(s-0)|,$$

de sorte qu'en réalité l'intégrale est envisagée seulement par rapport à l'ensemble $\{\varepsilon < |x| \leqslant \rho\}$ sur lequel la fonction $\varphi(x)$ est bornée.

Montrons que

$$\xi(t, A) = \int_A x\nu(t, dx). \tag{5}$$

En effet, si $A = \bigcup_{k=1}^{\infty} B_k$, où B_k sont des ensembles deux à deux disjoints dont les diamètres sont $\leqslant \delta$, et $x_k \in B_k$, alors

$$|\xi(t, A) - \sum_k x_k \nu(t, B_k)| \leqslant \sum_k |\xi(t, B_k) - x_k \nu(t, B_k)| \leqslant$$
$$\leqslant \delta \sum_k \nu(t, B_k) \leqslant \delta \nu(t, A)$$

($\xi(t, B_k)$ est la somme des $\nu(t, B_k)$ sauts appartenant à B_k et par conséquent distincts de x_k au plus de δ). D'où l'on déduit (5).

Comme

$$\mathsf{E}\, |\xi(t, A) - \sum_k x_k \nu(t, B_k)| \leqslant \delta \mathsf{E} \nu(t, A),$$
$$\mathsf{E}\, |\xi(t, A) - \sum_k x_k \nu(t, B_k)|^2 \leqslant \delta^2 \mathsf{E}\, [\nu(t, A)]^2,$$

il vient

$$\mathsf{E}\xi(t, A) = \lim_{\delta \to 0} \mathsf{E} \sum_k x_k \nu(t, B_k) = \int_A x\Pi(t, dx),$$
$$\mathrm{Var}(\xi(t, A), z) = \int_A (z, x)^2 \Pi(t, dx)$$

pour tout ensemble borné A situé à une distance positive de l'origine de l'espace X.

Soit la fonction aléatoire d'ensemble

$$\tilde{\nu}(t, A) = \nu(t, A) - \Pi(t, A).$$

Cette fonction possède les propriétés suivantes:

$$\mathsf{E}\tilde{\nu}(t, A) = 0,$$
$$\mathsf{E}(\tilde{\nu}(t, A)\, \tilde{\nu}(t, B)) = \mathsf{E}\nu(t, A \cap B) = \Pi(t, A \cap B). \tag{6}$$

Les égalités (6) permettent d'utiliser la construction générale de l'intégrale stochastique par rapport à une mesure orthogonale pour définir l'intégrale

$$\int f(x)\,\tilde{\nu}(t,\ dx)$$

$f(x)$ étant des fonctions mesurables telles que

$$\int |f(x)|^2\,\Pi(t,\ dx) < \infty$$

(cf. chapitre V, § 3). Au paragraphe précédent on a vu que

$$\sum_{k=2}^{\infty} \operatorname{Var}(\xi(t,\ \Delta_k),\ z) < \infty.$$

Comme

$$\operatorname{Var}(\xi(t,\ \Delta_k),\ z) = \int_{\Delta_k} (x,\ z)^2\,\Pi(t,\ dx),$$

pour tout $z \in X$

$$\lim_{\varepsilon\to 0} \int_{\varepsilon<|x|\leqslant\varepsilon_1} (x,\ z)^2\,\Pi(t,\ dx) < \infty.$$

Donc

$$\lim_{\varepsilon\to 0} \int_{\varepsilon<|x|\leqslant\varepsilon_1} |x|^2\,\Pi(t,\ dx) < \infty$$

et par suite existe

$$\int_{0<|x|<\varepsilon_1} x\tilde{\nu}(t,\ dx).$$

A noter d'autre part que

$$\sum_{k=2}^{n} (\xi(t,\ \Delta_k) - \mathsf{E}\xi(t,\ \Delta_k)) = \int_{\varepsilon_{n+1}<|x|\leqslant\varepsilon_1} x\tilde{\nu}(t,\ dx),$$

donc la série $\sum_{k=2}^{\infty} (\xi(t,\ \Delta_k) - \mathsf{E}\xi(t,\ \Delta_k))$ est stochastiquement convergente vers $\int_{0<|x|\leqslant\varepsilon_1} x\tilde{\nu}(t,\ dx)$. L'égalité

$$\xi(t,\ \Delta_1) = \int_{|x|>\varepsilon_1} x\nu(t,\ dx),$$

le théorème 2 et la remarque subséquente entraînent le théorème suivant (pour fixer les idées on posera $\varepsilon_1 = 1$).

THÉORÈME 3. *Si $\xi(t)$ est un processus stochastiquement continu séparable à accroissements indépendants, il existe presque sûrement un processus $\xi_0(t)$ continu à accroissements gaussiens indépendants ne dépendant pas de la mesure $\nu(t, A)$, tel que*

$$\xi(t) = \xi_0(t) + \int\limits_{|x|>1} x\nu(t, dx) + \int\limits_{|x|\leqslant 1} x\tilde{\nu}(t, dx). \tag{7}$$

Trouvons la fonction caractéristique de $\xi(t)$. Les termes du second membre de (7) étant indépendants les uns des autres, il suffit de trouver la fonction caractéristique de chacun d'eux. Celle de $\xi_0(t)$ est définie par la formule (1), § 3. Soit A ensemble borné de $\mathfrak{B}_\varepsilon$. On a

$$\mathsf{E}\exp\{i(z, \xi(t, A))\} = \lim_{\lambda\to 0} \mathsf{E}\exp\left\{i\left(z, \sum_k x_k\nu(t, B_k)\right)\right\},$$

où $A = \bigcup\limits_{k=1}^{n} B_k$, B_k étant des ensembles deux à deux disjoints de diamètres $\leqslant\lambda$, $x_k \in B_k$. Les quantités $\nu(t, B_k)$ sont mutuellement indépendantes et admettent une répartition poissonienne de paramètres $\Pi(t, B_k)$ respectivement. Donc

$$\mathsf{E}\exp\left\{i\left(z, \sum_k x_k\nu(t, B_k)\right)\right\} = \prod_{k=1}^{n} \mathsf{E}\exp\{i(z, x_k)\nu(t, B_k)\} =$$

$$= \prod_{k=1}^{n} \exp\{(e^{i(z, x_k)} - 1)\Pi(t, B_k)\} = \exp\left\{\sum_{k=1}^{n}(e^{i(z, x_k)} - 1)\Pi(t, B_k)\right\}.$$

En passant à la limite pour $\lambda \to 0$, on s'assure que

$$\mathsf{E}\exp\{i(z, \xi(t, A))\} = \exp\left\{\int\limits_A (e^{i(z, x)} - 1)\Pi(t, dx)\right\}. \tag{8}$$

La formule (8) est vraie pour des ensembles infinis de $\mathfrak{B}_\varepsilon$, puisqu'on peut les représenter par une somme monotone croissante de suites d'ensembles bornés A_n vérifiant (8) et passer à la limite sur n. De (8) il résulte

$$\mathsf{E}\exp\{i(z, \xi(t, A) - \mathsf{E}\xi(t, A))\} =$$
$$= \exp\left\{\int\limits_A (e^{i(z, x)} - 1 - i(z, x))\Pi(t, dx)\right\}, \tag{9}$$

i.e.

$$\mathsf{E}\exp\left\{i\left(z, \int\limits_A x\tilde{\nu}(t, dx)\right)\right\} =$$
$$= \exp\left\{\int\limits_A (e^{i(z, x)} - 1 - i(z, x))\Pi(t, dx)\right\}.$$

En passant une fois de plus à la limite sur A on s'assure que (9) est vraie pour tous les ensembles A qui en vérifient le second membre. En utilisant maintenant les formules (1), § 3, (8) et (9), on peut écrire la fonction caractéristique de $\xi(t)$:

$$\mathsf{E}e^{i(z,\,\xi(t))} = \mathsf{E}e^{i(z,\,\xi(0))} \exp\Big\{ i\,(a(t),\ z) - \frac{1}{2}(B(t)\,z,\ z) + \int\limits_{0<|x|>1} (e^{i(z,\,x)} - 1)\,\Pi(t,\ dx) + \int\limits_{0<|x|\leqslant 1} (e^{i(z,\,x)} - 1 - i\,(z,\ x))\,\Pi(t,\ dx)\Big\}. \quad (10)$$

Cette formule permet de déterminer la répartition de $\xi(t_2) - \xi(t_1)$ et partant toutes les répartitions finidimensionnelles du processus $\xi(t)$.

Puisque les processus stochastiquement équivalents admettent des répartitions finidimensionnelles égales et que tout processus est stochastiquement équivalent à un processus séparable (théorème 2, § 2, chapitre IV), on a le

THÉORÈME 4. *La fonction caractéristique d'un processus stochastiquement continu à accroissements indépendants est de la forme* (10), *où*

1) $\Pi(t,\ A)$ *est une fonction continue monotone non décroissante en* t *pour tout* $A \in \bigcup\limits_{\varepsilon} \mathfrak{B}_\varepsilon$ *et telle que*

$$\lim_{\varepsilon\to 0} \int\limits_{\varepsilon<|x|\leqslant 1} |x|^2\,\Pi(t,\ dx) < \infty\,;$$

2) $a(t)$ *est une fonction continue à valeurs dans* X ;

3) $B(t)$ *est une fonction continue dont les valeurs sont des opérateurs linéaires symétriques non négatifs de* X *et de plus* $B(t_2) - B(t_1)$ *sont aussi non négatifs pour* $t_1 < t_2$.

§ 5. Propriétés des réalisations

Dans ce paragraphe nous allons étudier quelques propriétés des réalisations de processus stochastiquement continus à accroissements indépendants. A noter qu'au § 2 on a trouvé les conditions nécessaires et suffisantes pour que les réalisations d'un processus soient en escalier et, au § 3, soient continues.

Soit $\xi(t)$ processus à valeurs numériques, c'est-à-dire X est la droite numérique. Etudions les conditions que doivent satisfaire les réalisations du processus $\xi(t)$ pour être presque sûrement des fonctions monotones.

THÉORÈME 1. *Pour que les réalisations de* $\xi(t)$, *processus stochastiquement continu, séparable, numérique, à accroissements indépen-*

dants, soient presque sûrement non décroissantes il faut et il suffit que la fonction caractéristique de $\xi(t)$ *soit représentable par*

$$\mathsf{E}e^{i\lambda\xi(t)} = \mathsf{E}e^{i\lambda\xi(0)} \exp\left\{i\lambda\gamma(t) + \int_0^\infty (e^{i\lambda x} - 1)\,\Pi(t, dx)\right\}, \qquad (1)$$

où la mesure Π *est telle que* $\int_0^1 x\Pi(t, dx) < \infty$, *et* $\gamma(t)$ *est une fonction non décroissante.*

D é m o n s t r a t i o n. *Condition nécessaire.* Si $\xi(t)$ est une fonction non décroissante, le processus $\xi(t)$ ne présente que des sauts positifs et par suite $\Pi(t, A) = 0$ pour tout ensemble $A \subset\,]-\infty, 0[$. On notera que le processus $\xi(t) - \xi(t, X_\varepsilon)$ sera aussi non décroissant (l'effacement des sauts ne viole pas la monotonie). Le processus $\xi(t, X_\varepsilon) - \xi(t, X_1)$, $0 < \varepsilon < 1$, qui est somme de sauts non négatifs et de plus

$$0 \leqslant \xi(t, X_\varepsilon) - \xi(t, X_1) \leqslant \xi(t) - \xi(0) - \xi(t, X_1),$$

sera aussi un processus monotone. Le lemme 1, § 4, indique que $\mathsf{E}\,[\xi(t) - \xi(0) - \xi(t, X_1)] < \infty$ et

$$\mathsf{E}\,[\xi(t, X_\varepsilon) - \xi(t, X_1)] = \int_\varepsilon^1 x\Pi(t, dx) \leqslant \mathsf{E}\,[\xi(t) - \xi(0) - \xi(t, X_1)]\,;$$

en passant à la limite pour $\varepsilon \to 0$, on vérifie que

$$\int_0^1 x\Pi(t, dx)$$

est finie. On remarquera par ailleurs que les $\xi(t) - \xi(t, X_\varepsilon)$ décroissent pour $\varepsilon \to 0$ (puisque de nouveaux sauts positifs sont effacés lorsque ε diminue). Donc $\lim\limits_{\varepsilon\to 0} [\xi(t) - \xi(t, X_\varepsilon)] = \xi_0(t)$ existe presque sûrement, le processus $\xi_0(t)$ étant presque sûrement continu. On a montré au § 3 que les accroissements du processus $\xi_0(t)$ admettaient des répartitions gaussiennes; or le processus $\xi_0(t)$ comme limite de processus non décroissants sera lui-même non décroissant, donc $\mathsf{P}\{\xi_0(t) - \xi_0(0) \geqslant 0\} = 1$. D'où il suit que $\operatorname{Var}[\xi_0(t) - \xi_0(0)] = 0$ (la variable ξ de répartition normale ne peut être presque sûrement non négative que pour $\operatorname{Var} \xi = 0$). Par suite

$$\xi_0(t) = \xi_0(0) + \gamma(t), \quad \text{où} \quad \gamma(t) = \mathsf{E}\,[\xi_0(t) - \xi_0(0)],$$

et par conséquent $\xi_0(t)$ ne décroît pas. La formule (1) peut être déduite de la relation

$$\mathsf{E}e^{i\lambda\xi(t)} = \lim_{\varepsilon\to\infty} \mathsf{E}e^{i\lambda\xi_0(t)}\mathsf{E}e^{i\lambda\xi(t, X_\varepsilon)}$$

si l'on tient compte de la forme du processus $\xi_0(t)$ et de la formule (8), § 4. Ceci achève la démonstration de la condition nécessaire.

Condition suffisante. Prouvons que $\mathsf{P}\{\xi(t_2) - \xi(t_1) \geqslant 0\} = 1$. Il suffit de montrer pour cela que la variable aléatoire ξ de fonction caractéristique

$$\mathsf{E}e^{i\lambda\xi} = \exp\left\{\int_0^\infty (e^{i\lambda x} - 1)\, dG(x)\right\}, \tag{2}$$

où $G(x)$ est une fonction monotone bornée, est presque sûrement $\geqslant 0$. Ceci découle du fait que la répartition de $\xi(t_2) - \xi(t_1)$ est la limite des répartitions de variables de fonction caractéristique

$$\exp\left\{i\lambda(\gamma(t_2) - \gamma(t_1)) + \int_\varepsilon^\infty (e^{i\lambda x} - 1)\, \Pi(t, dx)\right\}$$

pour $\varepsilon \downarrow 0$. Posons

$$F(x) = c\,[G(x) - G(+0)], \quad c = [G(+\infty) - G(+0)]^{-1}.$$

Alors

$$\mathsf{E}e^{i\lambda\xi} = \sum_{k=0}^\infty \frac{c^k}{k!} e^{-c}\left(\int_0^\infty e^{i\lambda x}\, dF(x)\right)^k,$$

de sorte que la fonction caractéristique de la variable ξ est confondue avec celle de S_ν, où $S_0 = 0$, $S_n = \xi_1 + \ldots + \xi_n$; $\xi_1, \xi_2, \ldots$ est une suite de variables non négatives indépendantes de même fonction de répartition $F(x)$, ν une variable aléatoire poissonienne ne dépendant pas de $\xi_1, \xi_2, \ldots$ Par suite $\xi \geqslant 0$.

Donc

$$\mathsf{P}\{\xi(t_2) \geqslant \xi(t_1)\} = 1 \text{ pour } t_1 < t_2.$$

D'où il résulte que l'événement : $\{\xi(t_1) \leqslant \xi(t_2)$ pour tous les couples de points rationnels tels que $t_1 < t_2\}$ a lieu presque sûrement. Si de plus l'on tient compte de ce que $\xi(t)$ ne présente presque sûrement pas de discontinuités de seconde espèce et est confondu presque sûrement soit avec $\xi(t-0)$, soit avec $\xi(t+0)$ (car le processus $\xi(t)$ est séparable), on obtient $\xi(t_1) \leqslant \xi(t_2)$ presque sûrement pour tous les couples t_1, t_2, tels que $t_1 < t_2$. ■

Etudions les conditions que doivent remplir les réalisations du processus $\xi(t)$ pour être presque sûrement à variation bornée.

On rappelle que la variation sur $[a, b]$ d'une fonction $x(t)$ définie sur $[a, b]$ et à valeurs dans X est la quantité

$$\operatorname*{var}_{[a,b]} x(t) = \sup \sum_{i=0}^{n-1} |x(t_i) - x(t_{i+1})|,$$

le suprémum étant pris sur toutes les partitions de l'intervalle $[a, b]$: $a = t_0 < t_1 < \ldots < t_n = b$.

THÉORÈME 2. *Pour que les réalisations de* $\xi(t)$, *processus stochastiquement continu séparable à accroissements indépendants défini sur* $[0, T]$, *soient presque sûrement à variation bornée sur* $[0, T]$, *il faut et il suffit que la fonction caractéristique de* $\xi(t)$ *soit donnée par la formule* (10), § 4, *où* $\operatorname{var}_{[a,b]} a(t) < \infty$, $B(t) = 0$ *et la mesure* $\Pi(t, A)$ *est telle que*

$$\int_{0<|x|\leq 1} |x| \, \Pi(t, dx) < \infty.$$

D é m o n s t r a t i o n. On commencera par la condition suffisante. Le processus défini par $\xi(t, X_1) = \int_{|x|>1} x\nu(t, dx)$ étant presque sûrement constant par morceaux, sa variation qui coïncide avec la somme des valeurs absolues des sauts sera finie. Par hypothèse la fonction $a(t)$ est à variation bornée. Donc, pour prouver que $\xi(t)$ l'est aussi, il suffit de prouver que l'intégrale $\int_{0<|x|\leq 1} x\tilde{\nu}(t, dx)$ le sera comme fonction de t (cf. formule (7), § 4). Soit le processus

$$\xi^{(\varepsilon)}(t) = \int_{\varepsilon<|x|\leq 1} x\tilde{\nu}(t, dx).$$

Il est immédiat de calculer

$$\operatorname{var}_{[0,T]} \xi^{(\varepsilon)}(t) = \operatorname{var}_{[0,T]} \int_{\varepsilon<|x|\leq 1} x\nu(t, dx) + \operatorname{var}_{[0,T]} \int_{\varepsilon<|x|\leq 1} x\Pi(t, dx).$$

Donc

$$\operatorname{var}_{[0,T]} \xi^{(\varepsilon)}(t) \leqslant \int_{\varepsilon<|x|\leq 1} |x| \, \nu(T, dx) +$$

$$+ \sup \sum_{k=1}^{n} \int_{\varepsilon<|x|\leq 1} |x| \, (\Pi(t_k, dx) - \Pi(t_{k-1}, dx)) =$$

$$= \int_{\varepsilon<|x|\leq 1} |x| \, \nu(T, dx) + \int_{\varepsilon<|x|\leq 1} |x| \, \Pi(T, dx) \leqslant$$

$$\lim_{\varepsilon\to 0} \int_{\varepsilon<|x|\leq 1} |x| \, \nu(T, dx) + \int_{0<|x|\leq 1} |x| \, \Pi(T, dx).$$

L'existence de $\lim\limits_{\varepsilon\to 0} \int\limits_{\varepsilon<|x|\leqslant 1} |x|\, \nu(T, dx)$ est une conséquence de ce que $\int\limits_{\varepsilon<|x|\leqslant 1} |x|\, \nu(T, dx)$ dépend de façon monotone de ε et

$$\mathsf{E} \int\limits_{\varepsilon<|x|\leqslant 1} |x|\, \nu(T, dx) \leqslant \int\limits_{0<|x|\leqslant 1} |x|\, \Pi(T, dx) < \infty.$$

On peut choisir une suite ε_n convergente vers zéro pour $n\to\infty$ telle que la suite de processus $\xi^{(\varepsilon_n)}(t)$ soit presque sûrement convergente vers le processus $\int\limits_{0<|x|\leqslant 1} x\tilde{\nu}(t, dx)$. On a donc presque sûrement

$$\operatorname*{var}_{[0,T]} \int\limits_{0<|x|\leqslant 1} x\tilde{\nu}(t, dx) \leqslant \int\limits_{0<|x|\leqslant 1} |x|\, \nu(T, dx) + \int\limits_{0<|x|\leqslant 1} |x|\, \Pi(T, dx).$$

(Il est aisé de voir que si $x_n(t)\to x(t)$ pour tout t, alors

$$\sum_{k=0}^{m-1} |x(t_{k+1}) - x(t_k)| = \lim_{n\to\infty} \sum_{k=0}^{m-1} |x_n(t_{k+1}) - x_n(t_k)| \leqslant \varlimsup_{n\to\infty} \operatorname*{var}_{[0,T]} x_n(t),$$

et par suite $\operatorname*{var}\limits_{[0,T]} x(t) \leqslant \varlimsup\limits_{n\to\infty} \operatorname*{var}\limits_{[0,T]} x_n(t)$.) Donc $\int\limits_{0<|x|\leqslant 1} x\tilde{\nu}(t, dx)$ est à variation bornée.

Prouvons maintenant la condition nécessaire. Supposons que $\xi(t)$ est à variation bornée sur $[0, T]$. Cette variation sera alors finie sur tout intervalle partiel de $[0, T]$. Soit le processus

$$\zeta(t) = \operatorname*{var}_{[0,t]} \xi(s).$$

Ce processus sera aussi à accroissements indépendants. Son accroissement $\zeta(t_2) - \zeta(t_1)$, $t_1 < t_2$, n'est autre que sa variation sur $[t_1, t_2]$ et ne dépend donc que des accroissements de $\xi(t)$ sur $[t_1, t_2]$. $\zeta(t)$ est stochastiquement continu, car $\operatorname*{var}\limits_{[0,t]} x(s)$ ne présente des discontinuités qu'aux points de discontinuité de $x(s)$ et le processus $\xi(t)$ ne possède pas de discontinuités fixes du fait de sa continuité stochastique. Il est évident enfin que $\zeta(t)$ est une fonction non décroissante. De ce qui précède il suit que

$$\zeta(t+0) - \zeta(t-0) = |\xi(t+0) - \xi(t-0)|.$$

Donc si $\zeta^{(\varepsilon)}(t)$ est la somme de tous les sauts de $\zeta(t)$ supérieurs à ε qui ont lieu sur $[0, t[$, alors

$$\zeta^{(\varepsilon)}(t) = \int\limits_{|x|>\varepsilon} |x|\, \nu(t, dx).$$

Le processus $\zeta(t)$ étant monotone, il vient

$$\zeta(t) = \gamma(t) + \lim_{\varepsilon \to 0} \zeta^{(\varepsilon)}(t) = \gamma(t) + \lim_{\varepsilon \to 0} \int_{|x| > \varepsilon} |x| \, \nu(t, dx)$$

(puisque $\zeta(0) = 0$), où $\gamma(t)$ est la variation de la composante continue du processus $\xi(t)$. D'autre part

$$\int_{|x| \leqslant 1} |x| \, \Pi(t, dx) = \mathsf{E}[\zeta(t) - \zeta^{(1)}(t)] < \infty.$$

A noter enfin que si $\xi_0(t)$ est la composante continue du processus $\xi(t)$, alors

$$|\xi_0(t) - \xi_0(0)| \leqslant \operatorname*{var}_{[0, T]} \xi_0(t) = \gamma(t)$$

et

$$|(\xi_0(t) - \xi_0(0), z)| \leqslant \gamma(t) |z|.$$

Or toute variable normalement répartie, de variance positive ne peut être presque sûrement bornée par aucune constante (indépendante du hasard). Donc pour tout z

$$\operatorname{Var}(z, \xi_0(t) - \xi_0(0)) = 0, \quad \text{i.e. } B(t) = 0.$$

Par suite $\gamma(t)$ est la variation de la fonction

$$a(t) = \int_{0 < |x| \leqslant 1} x \Pi(t, dx).$$

Comme

$$\operatorname*{var}_{[0, T]} \int_{0 < |x| \leqslant 1} x \Pi(t, dx) \leqslant \int_{0 < |x| \leqslant 1} |x| \, \Pi(T, dx)$$

est finie, il résulte que $\operatorname*{var}_{[0, T]} a(t) < \infty$. ■

REMARQUE. De la démonstration de la condition nécessaire il suit la proposition: *si un processus* $\xi(t)$ *est défini par*

$$\xi(t) = \xi(0) + a_1(t) + \int_{|x| > 0} x \nu(t, dx),$$

alors

$$\zeta(t) = \operatorname*{var}_{[0, t]} \xi(s) = \operatorname*{var}_{[0, t]} a_1(s) + \int_{|x| > 0} |x| \, \nu(t, dx)$$

et la fonction caractéristique de $\zeta(t)$ *est donnée par*

$$\mathsf{E} e^{i\lambda\zeta(t)} = \exp\left\{ i\lambda \operatorname*{var}_{[0, t]} a_1(s) + \int_{|x| > 0} (e^{i\lambda|x|} - 1) \Pi(t, dx) \right\}.$$

Etudions maintenant un processus homogène à valeurs dans $\mathcal{R}^1$. On s'intéressera au comportement de $\xi(t)$ pour $t \downarrow 0$ et $t \to \infty$.

Toute fonction $g(t)$ continue et monotone croissante pour $t > 0$ sera dite à *croissance régulière* si existent des fonctions $k_1(\lambda)$ et $k_2(\lambda)$ telles que $k_1(\lambda) \to \infty$ avec λ et $k_2(\lambda) \to 1$ avec λ et

$$k_1(\lambda)\, g(t) \leqslant g(\lambda t) \leqslant k_2(\lambda)\, g(t).$$

Une fonction à croissance régulière $g(t)$ est dite *fonction supérieure* d'un processus $\xi(t)$ pour $t \uparrow \infty$ (resp. $t \downarrow 0$) si

$$\mathsf{P}\left\{\overline{\lim_{t\to\infty}} \frac{\xi(t)}{g(t)} > 1\right\} = 1 \quad \left(\text{resp. } \mathsf{P}\left\{\overline{\lim_{t\downarrow 0}} \frac{\xi(t)}{g(t)} > 1\right\} = 1\right),$$

et *fonction inférieure* pour $t \uparrow \infty$ (resp. $t \downarrow 0$) si

$$\mathsf{P}\left\{\overline{\lim_{t\to\infty}} \frac{\xi(t)}{g(t)} < 1\right\} = 1 \quad \left(\text{resp. } \mathsf{P}\left\{\overline{\lim_{t\downarrow 0}} \frac{\xi(t)}{g(t)} < 1\right\} = 1\right).$$

Etudions tout d'abord les fonctions supérieures et inférieures de processus homogènes symétriques à accroissements indépendants et ensuite celles de $|\xi(t)|$, où $\xi(t)$ n'est pas forcément un processus symétrique.

On rappelle qu'un processus $\xi(t)$ est *symétrique* s'il a les mêmes répartitions finidimensionnelles que $-\xi(t)$. On aura besoin du

LEMME 1. *Si $\xi(t)$ est un processus stochastiquement continu séparable symétrique à accroissements indépendants, alors*

$$\mathsf{P}\{\sup_{0\leqslant t\leqslant T} \xi(t) > c\} \leqslant 2\mathsf{P}\{\xi(T) > c\}. \tag{3}$$

Démonstration. Soient $\xi_1, \xi_2, \ldots, \xi_n$ des variables aléatoires indépendantes et symétriquement réparties, $S_k = \xi_1 + \ldots + \xi_k$. On a

$$\mathsf{P}\{\sup_k S_k > c\} \leqslant 2\mathsf{P}\{S_n > c\}. \tag{4}$$

En effet,

$$2\mathsf{P}\{S_n - S_k \geqslant 0\} \geqslant 1$$

entraîne

$$\mathsf{P}\{\sup_k S_k > c\} = \sum_{k=1}^{n} \mathsf{P}\{\sup_{i\leqslant k-1} S_i \leqslant c,\ S_k > c\} \leqslant$$

$$\leqslant \sum_{k=1}^{n} \mathsf{P}\{\sup_{i\leqslant k-1} S_i \leqslant c,\ S_k > c\}\, 2\mathsf{P}\{S_n - S_k \geqslant 0\} =$$

$$= 2\sum_{k=1}^{n} \mathsf{P}\{\sup_{i\leqslant k-1} S_i \leqslant c,\ S_k > c,\ S_n - S_k \geqslant 0\}.$$

Les événements $\{\sup_{i\leqslant k-1} S_i\leqslant c,\ S_k>c,\ S_n-S_k\geqslant 0\}$ sont incompatibles et chacun d'eux implique l'événement $\{S_n>c\}$. Donc

$$2\mathsf{P}\{S_n>c\}\geqslant 2\sum_{k=1}^{n}\mathsf{P}\{\sup_{i\leqslant k-1} S_i\leqslant c, S_k>c, S_n-S_k\geqslant 0\}\geqslant$$
$$\geqslant \mathsf{P}\{\sup_k S_k>c\}.$$

La formule (4) est établie. Appliquons-la maintenant à

$$\xi_k=\xi\left(\frac{k}{n}T\right)-\xi\left(\frac{k-1}{n}T\right),\ \xi_1=\xi\left(\frac{T}{n}\right).$$

On trouve

$$\mathsf{P}\left\{\sup_{1\leqslant k\leqslant n}\xi\left(\frac{kT}{n}\right)>c\right\}\leqslant 2\mathsf{P}\{\xi(T)>c\}.$$

En passant à la limite pour $n\to\infty$ et comme $\xi(t)$ ne présente presque sûrement pas de discontinuités de seconde espèce de sorte que

$$\mathsf{P}\left\{\lim_{n\to\infty}\sup_{1\leqslant k\leqslant n}\xi\left(\frac{kT}{n}\right)=\sup_{0\leqslant t\leqslant T}\xi(t)\right\}=1,$$

on obtient (3). ■

Théorème 3. *Si $\xi(t)$ est un processus stochastiquement continu symétrique homogène séparable et à accroissements indépendants, $g(t)$ une fonction à croissance régulière telle que*

$$\int_1^\infty \frac{1}{t}\mathsf{P}\{\xi(t)>g(t)\}\,dt<\infty,$$

alors pour tout $\lambda>1$ la fonction $\lambda g(t)$ sera fonction supérieure pour le processus $\xi(t)$ pour $t\to\infty$.

Démonstration. Soit $a>1$. Appelons A_k l'événement $\{\sup_{0\leqslant t\leqslant a^k}\xi(t)>g(a^{k+1})\}$. Le lemme (1) nous dit que

$$\mathsf{P}\{A_k\}\leqslant 2\mathsf{P}\{\xi(a^k)>g(a^{k+1})\}.$$

Or pour $t\in[a^k, a^{k+1}]$ on aura

$$g(t)<g(a^{k+1}),\ \ 2\mathsf{P}\{\xi(t)-\xi(a^k)\geqslant 0\}\geqslant 1,$$

donc

$$\mathsf{P}\{A_k\}\leqslant 4\mathsf{P}\{\xi(a^k)>g(t)\}\,\mathsf{P}\{\xi(t)-\xi(a^k)\geqslant 0\}\leqslant 4\mathsf{P}\{\xi(t)>g(t)\}.$$

Par suite

$$\int_{a^k}^{a^{k+1}} \frac{1}{t} \mathsf{P}\{A_k\}\, dt \leqslant 4 \int_{a^k}^{a^{k+1}} \frac{1}{t} \mathsf{P}\{\xi(t) > g(t)\}\, dt$$

et

$$\sum_{k=1}^{n} \mathsf{P}\{A_k\} \leqslant \frac{4}{\ln a} \int_{1}^{\infty} \frac{1}{t} \mathsf{P}\{\xi(t) > g(t)\}\, dt.$$

Le lemme de Borel-Cantelli nous apprend que seul un nombre fini d'événements A_k a presque sûrement lieu, c'est-à-dire à partir d'un certain rang (en général aléatoire) les événements A_k n'ont pas lieu. Donc

$$\mathsf{P}\left\{\overline{\lim_{k\to\infty}} \frac{1}{g(a^{k+1})} \sup_{a^{k-1} \leqslant t \leqslant a^k} \xi(t) \leqslant 1\right\} = 1.$$

Pour $t \in [a^{k-1}, a^k]$ on a

$$\frac{\xi(t)}{k_2(a^2)\, g(t)} \leqslant \frac{1}{k_2(a^2)\, g(a^{k-1})} \sup_{a^{k-1} \leqslant t \leqslant a^k} \xi(t) \leqslant \frac{1}{g(a^{k+1})} \sup_{a^{k-1} \leqslant t \leqslant a^k} \xi(t).$$

Par suite $\lambda g(t)$ sera fonction supérieure pour $\xi(t)$ quels que soient $a > 1$ et $\lambda > k_2(a^2)$. ■

Le théorème est visiblement vrai pour un processus $\xi(t)$ non homogène, puisque cette propriété n'est pas intervenue dans la démonstration.

THÉORÈME 4. *Si $\xi(t)$ est un processus homogène symétrique à accroissements indépendants et $g(t)$ une fonction à croissance régulière telle que la série*

$$\sum_{k=1}^{\infty} \mathsf{P}\{\xi(a^k) > g(a^k)\}$$

diverge pour tout $a > 1$, alors $\lambda g(t)$ sera fonction inférieure pour le processus $\xi(t)$, $\forall \lambda \in]0, 1[$ pour $t \to \infty$.

Démonstration. On distinguera deux cas.

1)

$$\underline{\lim_{t\to\infty}} \mathsf{P}\{|\xi(t)| < g(t)\} = 0.$$

Il existe alors une suite $t_k \to \infty$ telle que

$$\mathsf{P}\{|\xi(t_k)| < g(t_k)\} < 2^{-k}.$$

Donc, en vertu du lemme de Borel-Cantelli

$$\mathsf{P}\left\{\overline{\lim_{k\to\infty}} \frac{|\xi(t_k)|}{g(t_k)} \geqslant 1\right\} = 1.$$

Et pour tout $\lambda \in]0, 1[$

$$\mathsf{P}\left\{\overline{\lim_{t\to\infty}} \frac{|\xi(t)|}{\lambda g(t)} > 1\right\} = 1. \tag{5}$$

2)

$$\underline{\lim_{t\to\infty}} \mathsf{P}\{|\xi(t)| < g(t)\} > \delta > 0.$$

Pour t assez grand on a alors

$$\mathsf{P}\{|\xi(t)| < g(t)\} > \delta.$$

Soient les événements indépendants

$$B_k = \{\xi(a^{k+1}) - \xi(a^k) > g(a^{k+1}) - g(a^k)\},$$

où $a > 1$. Il vient

$$\mathsf{P}(B_k) \geqslant \int_{-g(a^k)}^{g(a^k)} \mathsf{P}\{\xi(a^{k+1}) - z > g(a^{k+1}) - g(a^k)\} \mathsf{P}\{\xi(a^k) \in dz\} \geqslant$$

$$\geqslant \mathsf{P}\{\xi(a^{k+1}) > g(a^{k+1})\} \int_{-g(a^k)}^{g(a^k)} \mathsf{P}\{\xi(a^k) \in dz\} =$$

$$= \mathsf{P}\{\xi(a^{k+1}) > g(a^{k+1})\} \mathsf{P}\{|\xi(a^k)| \leqslant g(a^k)\} \geqslant \mathsf{P}\{\xi(a^{k+1}) > g(a^{k+1})\}\,\delta$$

pour k assez grand.

Donc la série $\sum_{k=1}^{\infty} \mathsf{P}\{B_k\}$ est divergente et en vertu du lemme de Borel-Cantelli une infinité d'événements B_k seront presque sûrement réalisés. A noter que l'événement B_k implique l'un des événements

$$\{-\xi(a^k) > g(a^k)\},\ \{\xi(a^{k+1}) > g(a^{k+1}) - 2g(a^k)\}.$$

Par suite une infinité d'événements

$$\{|\xi(a^k)| > g(a^k) - 2g(a^{k-1})\}$$

ont presque sûrement lieu. En choisissant a tel que

$$g(a^k) - 2g(a^{k-1}) = g(a^k)\left[1 - \frac{2g(a^{k-1})}{g(a^k)}\right] \geqslant$$

$$\geqslant g(a^k)\left[1 - \frac{2}{k_1(a)}\right] > \lambda g(a^k) \quad (0 < \lambda < 1)$$

(ce choix est possible, car $g(t)$ est à croissance régulière), on s'assure que (5) est réalisée.

Soient les événements

$$C = \left\{\overline{\lim_{t\to\infty}} \frac{\xi(t)}{\lambda g(t)} > 1\right\}, \quad D = \left\{\overline{\lim_{t\to\infty}} \frac{-\xi(t)}{\lambda g(t)} > 1\right\}.$$

De (5) il résulte que $\mathsf{P}(C \cup D) = 1$. La symétrie du processus $\xi(t)$ donne $\mathsf{P}(C) = \mathsf{P}(D)$. Enfin, la loi de tout ou rien (théorème 7, § 4, chapitre 2) nous apprend que $\mathsf{P}(C)$ et $\mathsf{P}(D)$ sont égales soit à l'unité, soit à zéro. Donc $\mathsf{P}(C) = \mathsf{P}(D) = 1$, sinon on aurait $\mathsf{P}(C) = 0$ et partant $\mathsf{P}(C \cup D) = 0$, ce qui contredit (5). ■

Remarque 1. Si l'on prend $a < 1$, sans rien changer à la démonstration des théorèmes 3 et 4, on s'assure de la validité des propositions suivantes:

soit $\varphi(t)$ *une fonction à croissance régulière;*

a) *si* $\xi(t)$ *est un processus stochastiquement continu symétrique séparable homogène à accroissements indépendants,*

$$\int_0^1 \frac{1}{t}\, \mathsf{P}\{\xi(t) > \varphi(t)\}\, dt < \infty,$$

alors pour $\lambda > 1$ *la fonction* $\lambda\varphi(t)$ *sera supérieure pour* $\xi(t)$ *pour* $t \downarrow 0$, *i.e.*

$$\mathsf{P}\left\{\overline{\lim_{t\downarrow 0}}\, \frac{\xi(t)}{\lambda\varphi(t)} < 1\right\} = 1;$$

b) *si* $\xi(t)$ *est un processus symétrique homogène à accroissements indépendants tel que la série* $\sum_{k=1}^{\infty} \mathsf{P}\{\xi(a^k) > \varphi(a^k)\}$ *diverge* $\forall a < 1$, *alors pour* $\lambda < 1$ *la fonction* $\lambda\varphi(t)$ *sera inférieure pour* $\xi(t)$ *pour* $t \downarrow 0$, *i.e.*

$$\mathsf{P}\left\{\overline{\lim_{t\to 0}}\, \frac{\xi(t)}{\lambda\varphi(t)} > 1\right\} = 1.$$

Appliquons les résultats obtenus au processus du mouvement brownien. Ce processus est symétrique.

Les majorations

$$\frac{1}{\sqrt{2\pi}} \int_z^{\infty} e^{-u^2/2}\, du \leqslant \frac{1}{\sqrt{2\pi}} \int_z^{\infty} \frac{u}{z}\, e^{-u^2/2}\, du = \frac{1}{\sqrt{2\pi}\, z}\, e^{-z^2/2} \; (z > 0),$$

$$\frac{1}{\sqrt{2\pi}} \int_z^{\infty} e^{-u^2/2}\, du \geqslant \frac{1}{\sqrt{2\pi}} \int_z^{z+1} e^{-u^2/2}\, du \geqslant \frac{1}{\sqrt{2\pi}}\, e^{-(z+1)^2/2} \; (z > 0)$$

entraînent

$$\frac{1}{\sqrt{2\pi}} \exp\left\{-\left(\frac{z}{\sqrt{t}} + 1\right)^2 \cdot \frac{1}{2}\right\} \leqslant \mathsf{P}\{w(t) > z\} \leqslant \sqrt{\frac{t}{2\pi}}\, \frac{1}{z}\, e^{-\frac{z^2}{2t}}.$$

Montrons que les fonctions $(1+\varepsilon)\sqrt{2t\ln\ln t}$ et $(1-\varepsilon)\sqrt{2t\ln\ln t}$ seront respectivement supérieure et inférieure $\forall\varepsilon\in\,]0,\,1[$. En effet,

$$\mathsf{P}\{w(t)>(1+\varepsilon)\sqrt{2t\ln\ln t}\}\leqslant$$
$$\leqslant\frac{1}{\sqrt{2\pi(1+\varepsilon)^2 2\ln\ln t}}\exp\left\{-\frac{(1+\varepsilon)^2 2\ln\ln t}{2}\right\}=O((\ln t)^{-(1+\varepsilon)^2}),$$

et l'intégrale $\int\limits_c^\infty \frac{dt}{t(\ln t)^{(1+\varepsilon)^2}}$ est convergente pour $c>1$. D'autre part,

$$\mathsf{P}\{w(a^k)>(1-\varepsilon)\sqrt{2a^k\ln\ln a^k}\}\geqslant$$
$$\geqslant\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}((1-\varepsilon)\sqrt{2\ln\ln a^k}+1)^2\right\}=$$
$$=\frac{e^{-1/2}}{\sqrt{2\pi}}\exp\{-(1-\varepsilon)^2\ln\ln a^k-(1-\varepsilon)\sqrt{2\ln\ln a^k}\}\geqslant$$
$$\geqslant\frac{e^{-1/2}}{\sqrt{2\pi}}[\ln a^k]^{-\alpha}=\frac{c_1}{k^\alpha},$$

où c_1 est une constante, pourvu que

$$\alpha>(1-\varepsilon)^2+(1-\varepsilon)\sqrt{\frac{2}{\ln\ln a^k}}.$$

Il est évident que l'on peut prendre $\alpha<1$ pour k assez grand. Donc la série $\sum \mathsf{P}\{w(a^k)>(1-\varepsilon)\sqrt{2a^k\ln\ln a^k}\}$ est divergente. Ainsi on a démontré le

THÉORÈME 5. *Si $w(t)$ est un processus du mouvement brownien séparable, alors*

$$\mathsf{P}\left\{\overline{\lim_{t\to\infty}}\frac{w(t)}{\sqrt{2t\ln\ln t}}=1\right\}=1.$$

En utilisant la remarque 1 on établit le

THÉORÈME 6. *Si $w(t)$ est un processus du mouvement brownien séparable, alors*

$$\mathsf{P}\left\{\overline{\lim_{t\to 0}}\frac{w(t)}{\sqrt{2t\ln\ln\frac{1}{t}}}=1\right\}=1.$$

Les théorèmes 5 et 6 s'appellent *loi du logarithme itéré.*

Les propositions auxiliaires suivantes nous seront utiles pour étudier les fonctions supérieure et inférieure pour $|\xi(t)|$, où $\xi(t)$ est un processus à accroissements indépendants.

LEMME 2. *Soient* $\xi_1, \ldots, \xi_n$ *des variables aléatoires indépendantes,* $S_k = \xi_1 + \ldots + \xi_k$ *et*

$$\mathsf{P}\{|S_n - S_k| > c\} \leqslant \alpha, \quad k = 1, 2, \ldots, n,$$

pour certains $\alpha < 1$ *et* $c > 0$. *Alors pour* $a > 0$

$$\mathsf{P}\{\sup_k |S_k| > a + c\} \leqslant \frac{1}{1-\alpha} \mathsf{P}\{|S_n| > a\}.$$

D é m o n s t r a t i o n. On a

$$\mathsf{P}\left(\bigcup_k [\{\sup_{i \leqslant k-1} |S_i| \leqslant a + c\} \cap \{|S_k| > a + c\} \cap \leqslant \{|S_n - S_k| \leqslant c\}]\right) \leqslant$$
$$\leqslant \mathsf{P}\{|S_n| > a\}$$

Donc

$$\sum_k \mathsf{P}(\{\sup_{i \leqslant k-1} |S_i| \leqslant a + c\} \cap \{|S_k| > a + c\}) \mathsf{P}\{|S_n - S_k| \leqslant c\} \leqslant$$
$$\leqslant \mathsf{P}\{|S_n| > a\},$$

$$\left[\sum_k \mathsf{P}(\{\sup_{i \leqslant k-1} |S_i| \leqslant a + c\} \cap \{|S_k| > a + c\})\right](1 - \alpha) \leqslant \mathsf{P}\{|S_n| > a\}.$$

Entre crochets on reconnaît $\mathsf{P}\{\sup_k |S_k| > a + c\}$. ■

LEMME 3. *Si* $\xi(t)$ *est un processus stochastiquement continu séparable à accroissements indépendants pour lequel existe* $\alpha < 1$ *tel que*

$$\mathsf{P}\{|\xi(T) - \xi(s)| > c\} \leqslant \alpha$$

pour $0 < s \leqslant T$, *alors*

$$\mathsf{P}\{\sup_{0 \leqslant s \leqslant T} |\xi(s)| > c + x\} \leqslant \frac{1}{1-\alpha} \mathsf{P}\{|\xi(T)| > x\}, \ \forall x > 0. \tag{6}$$

D é m o n s t r a t i o n. Le lemme 2 entraîne

$$\mathsf{P}\left\{\sup_{1 \leqslant k \leqslant n} \left|\xi\left(\frac{k}{n} T\right)\right| > x + c\right\} \leqslant \frac{1}{1-\alpha} \mathsf{P}\{|\xi(T)| > x\}.$$

En passant à la limite pour $n \to \infty$ on obtient ce qu'on voulait. ■

THÉORÈME 7. *Si* $\xi(t)$ *est un processus stochastiquement continu séparable homogène à accroissements continus et* $g(t)$ *une fonction à croissance régulière, telle que pour tout* $\varepsilon > 0$

$$\overline{\lim_{t \to \infty}} \mathsf{P}\{|\xi(t)| > \varepsilon g(t)\} < 1$$

et

$$\int_1^\infty \frac{1}{t} \mathsf{P}\{|\xi(t)| > g(t)\} \, dt < \infty,$$

alors pour $\lambda > 1$ *la fonction* $\lambda g(t)$ *sera fonction supérieure pour* $|\xi(t)|$, *i.e.*

$$\mathsf{P}\left\{\overline{\lim_{t\to\infty}}\frac{|\xi(t)|}{g(t)}\leqslant 1\right\}=1.$$

Démonstration. Soient $a\in]1, 2[$ et $\varepsilon>0$. Appelons A_k l'événement $\{\sup\limits_{t\leqslant a^k}|\xi(t)|>(1+2\varepsilon)g(a^{k+1})\}$. D'après les conditions du théorème il existe $c>0$ et T_0 tels que $\mathsf{P}\{|\xi(t)|>\varepsilon g(t)\}\leqslant 1-c$ pour $t>T_0$. Donc pour k assez grand et $t<a^k$ on a

$$\mathsf{P}\{|\xi(a^k)-\xi(t)|>\varepsilon g(^{k+1})\}=\mathsf{P}\{|\xi(a^k-t)|>\varepsilon g(a^{k+1})\}\leqslant$$
$$\leqslant\left\{\begin{array}{l}\sup\limits_{s\geqslant T_0}\mathsf{P}\{|\xi(s)|>\varepsilon g(s)\}\\ \sup\limits_{s\leqslant T_0}\mathsf{P}\{|\xi(s)|>\varepsilon g(a^{k+1})\}\end{array}\right\}\leqslant 1-c,$$

et en vertu du lemme 3

$$\mathsf{P}\{A_k\}\leqslant\frac{1}{c}\mathsf{P}\{|\xi(a^k)|>(1+\varepsilon)g(a^{k+1})\}.$$

Pour $t\in[a^k, a^{k+1}]$ on a

$$t-a^k\leqslant(a-1)a^k\leqslant a^k$$

et

$$\mathsf{P}\{|\xi(t)-\xi(a^k)|>\varepsilon g(a^k)\}<1-c,$$

et par suite pour k assez grand

$$\mathsf{P}\{A_k\}\leqslant\frac{1}{c^2}\mathsf{P}\{|\xi(a^k)|>(1+\varepsilon)g(a^{k+1})\}\times$$
$$\times\mathsf{P}\{|\xi(t)-\xi(a^k)|\leqslant\varepsilon g(a^{k+1})\}\leqslant\frac{1}{c^2}\mathsf{P}\{|\xi(t)|>g(a^{k+1})\}\leqslant$$
$$\leqslant\frac{1}{c^2}\mathsf{P}\{|\xi(t)|>g(t)\},$$

de sorte que

$$\mathsf{P}\{A_k\}\leqslant\frac{(\ln a)^{-1}}{c^2}\int\limits_{a^k}^{a^{k+1}}\frac{1}{t}\mathsf{P}\{|\xi(t)|>g(t)\}\,dt,$$

$$\sum_{k=l}^{\infty}\mathsf{P}\{A_k\}\leqslant(\ln a)^{-1}c^{-2}\int\limits_{a^l}^{\infty}\frac{1}{t}\mathsf{P}\{|\xi(t)|>g(t)\}\,dt.$$

Donc, seul un nombre fini d'événements A_k a presque sûrement lieu. En reprenant les raisonnements du théorème 4, on s'assure que

$\lambda(1 + 2\varepsilon) k_2(a^2) g(t)$ pour $\lambda > 1$ sera fonction supérieure pour $|\xi(t)|$. Le théorème résulte de ce que $k_2(a) \to 1$ pour $a \to 1$ et du fait que $a - 1$ et $\varepsilon > 0$ peuvent être rendus arbitrairement petits. ■

En analysant la démonstration du théorème 4 on déduit sans peine le

THÉORÈME 8. *Si $\xi(t)$ est un processus homogène à accroissements indépendants et $g(t)$ une fonction à croissance régulière telle que pour $a > 1$ la série*

$$\sum_{k=1}^{\infty} \mathsf{P}\{|\xi(a^k)| > g(a^k)\}$$

est divergente, alors la fonction $\lambda g(t)$ sera fonction inférieure pour $|\xi(t)|$, $\forall \lambda \in]0, 1[$, i.e.

$$\mathsf{P}\left\{\varlimsup_{t\to\infty} \frac{|\xi(t)|}{g(t)} \geqslant 1\right\} = 1.$$

Les théorèmes 7 et 8 peuvent être énoncés pour le cas où $t \to 0$, comme nous l'avons fait pour la remarque 1.

CHAPITRE VII

PROCESSUS MARKOVIENS DISCONTINUS

§ 1. Définition générale du processus markovien

Les processus markoviens au sens large ont été étudiés au § 4, chapitre I. On rappelle qu'on y a introduit la probabilité $\mathsf{P}(s, x, t, A)$ que le processus se retrouve à l'instant t dans un ensemble d'états A sachant qu'à l'instant s il était dans l'état x. Les probabilités de passage permettent grâce au théorème de Kolmogorov (chapitre III, § 2) de construire toute une famille de mesures probabilistes sur l'espace des fonctions $\mathsf{P}_{s,x}$ définies sur les ensembles cylindriques de l'espace $F_{[s,\infty[}$ de toutes les fonctions données sur $[s, \infty[$ par

$$\mathsf{P}_{s,x}\{x(\cdot): x(t_1) \in A_1, \ldots, x(t_k) \in A_k\} = \\ = \int_{A_1} \mathsf{P}(s, x, t_1, dx_1) \ldots \int_{A_k} \mathsf{P}(t_{k-1}, x_{k-1}, t_k, dx_k) \quad (1)$$

pour $s < t_1 < \ldots < t_k$. Les mesures $\mathsf{P}_{s,x}$ seront traitées naturellement comme les répartitions conditionnelles d'un processus markovien étudié sur l'intervalle $[s, \infty[$ sachant qu'à l'instant s il se trouvait dans l'état x. Contrairement donc à la définition du processus aléatoire ordinaire, qui ne fait intervenir qu'un espace probabilisé, on fera l'étude des processus markoviens en se basant sur une famille d'espaces probabilisés différant par leurs tribus et leurs mesures.

Au chapitre IV on a montré que dans de nombreux cas il était possible de construire une mesure adaptée aux répartitions finidimensionnelles du processus sur un espace fonctionnel plus étroit que celui de toutes les fonctions. Les processus dont les réalisations possèdent certaines propriétés de régularité constituent une riche matière d'étude et sont très importants dans les applications. Aussi supposerons-nous que les réalisations d'un processus assujetti à démarrer à l'instant s, appartiennent à un espace fonctionnel F_s ($s \geqslant 0$). Les espaces fonctionnels F_s sont adaptés de la manière suivante : la restriction de toute fonction $x(\cdot) \in F_s$ à $[u, \infty[$ ($u > s$) appartient à F_u.

A noter enfin que la notion de « passé » est essentielle dans l'étude des processus markoviens. Si un processus débute à l'instant s, son « passé » à l'instant t est défini par les événements observables

sur l'intervalle $[s, t]$. Appelons $\mathfrak{S}_t^s$ l'ensemble de ces événements. On supposera naturellement que $\mathfrak{S}_t^s$ est une tribu, $\mathfrak{S}_{t_1}^s \subset \mathfrak{S}_{t_2}^s$ pour $t_1 < t_2$ et que la valeur du processus à l'instant t est mesurable par rapport à $\mathfrak{S}_t^s$ (autrement dit, la valeur du processus est observable à cet instant).

Récapitulons. Un processus markovien sera donc défini par la donnée des éléments suivants:

a) $\{X, \mathfrak{B}\}$, espace probabilisable et espace des phases du processus;

b) F_s $(s \geqslant 0)$, famille d'espaces de fonctions à valeurs dans X vérifiant la condition de compatibilité: la restriction de toute fonction de F_s à $[u, \infty[$ appartient à F_u pour $u > s$;

c) $\{\Omega, \mathfrak{S}\}$, espace probabilisable et ensemble des tribus $\mathfrak{S}_t^s \subset \mathfrak{S}$ définies pour $0 \leqslant s \leqslant t < \infty$ telles que $\mathfrak{S}_{t_1}^{s_1} \subset \mathfrak{S}_{t_2}^{s_2}$ pour $[s_1, t_1] \subset [s_2, t_2]$;

d) une famille de mesures $\mathsf{P}_{s,x}$, $s \in [0, \infty[$, $x \in X$, définie sur les tribus $\mathfrak{S}^s = \bigcup\limits_t \mathfrak{S}_t^s$;

e) $x_s(t, \omega)$, fonction définie sur $[s, \infty[$ et telle que:

1) $x_s(\cdot, \omega) \in F_s$ pour $\omega \in \Omega$;

2) $x_s(t, \omega)$ est mesurable par rapport à $\mathfrak{S}_t^u$ $\forall u \in [s, t]$;

3) $\mathsf{P}_{s,x}(\{\omega : x_s(s, \omega) = x\}) = 1$;

4) $\mathsf{P}_{s,x}(\{\omega : x_s(t, \omega) \in B\} \mid \mathfrak{S}_u^s) =$
$$= \mathsf{P}_{u, x_s(u,\omega)}(\{\omega : x_u(t, \omega) \in B\}) \pmod{\mathsf{P}_{s,x}} \qquad (2)$$
pour $B \in \mathfrak{B}$, $s < u < t$.

Appelons $\{x_s(t, \omega), \mathfrak{S}_t^s, \mathsf{P}_{s,x}\}$ le processus markovien défini avec les éléments énumérés; $x_s(\cdot, \omega)$ ses réalisations; $\mathfrak{S}_t^s$ la tribu des observables sur $[s, t]$; $\mathsf{P}_{s,x}$ la répartition des probabilités du processus issu de x à l'instant s; (2) définit la propriété fondamentale d'être markovien, c'est-à-dire l'indépendance par rapport au passé, le présent étant donné.

Les probabilités de passage se définissent maintenant par
$$\mathsf{P}(s, x, t, B) = \mathsf{P}_{s,x}(\{\omega : x_s(t, \omega) \in B\}). \qquad (3)$$

On admettra qu'est toujours réalisée la condition

5) la probabilité de passage est mesurable en x par rapport à $\mathfrak{B}$.

De (2) on déduit aussitôt l'équation de Kolmogorov-Chapman:
$$\begin{aligned}\mathsf{P}(s, x, t, B) &= \mathsf{P}_{s,x}(\{\omega : x_s(t, \omega) \in B\}) = \\ &= \mathsf{E}_{s,x}\mathsf{P}_{s,x}(\{\omega : x_s(t, \omega) \in B\} \mid \mathfrak{S}_u^s) = \\ &= \mathsf{E}_{s,x}\mathsf{P}_{u, x_s(u,\omega)}(\{\omega : x_u(t, \omega) \in B\}) = \\ &= \mathsf{E}_{s,x}\mathsf{P}(u, x_s(u, \omega), t, B) = \\ &= \int \mathsf{P}(s, x, u, dy)\,\mathsf{P}(u, y, t, B)\end{aligned}$$

pour $s < u < t$. Ici et dans la suite $\mathsf{E}_{s,x}$ désignera l'espérance mathématique par rapport à la mesure $\mathsf{P}_{s,x}$ d'une variable aléatoire $\xi(\omega)$ $\mathfrak{S}^s$-mesurable

$$\mathsf{E}_{s,x}\xi(\omega) = \int \xi(\omega)\, \mathsf{P}_{s,x}(d\omega). \tag{4}$$

L'expression (2) peut être transcrite en termes d'espérance mathématique:

$$\mathsf{E}_{s,x}(g(x_s(t,\ \omega)) \mid \mathfrak{S}_u^s) = \mathsf{E}_{u,\, x_s(u,\, \omega)} g(x_u(t,\ \omega)) \pmod{\mathsf{P}_{s,x}}, \tag{5}$$

$s < u < t$, pour toute fonction g $\mathfrak{B}$-mesurable bornée. En effet, si $g(x) = \chi_B(x)$, où χ_B est l'indicateur de B, (2) découle de (5). Par ailleurs, (2) entraîne (5) pour les indicateurs d'ensembles mesurables, et partant pour les fonctions simples, combinaisons linéaires d'indicateurs. Reste à noter que toute fonction mesurable bornée est limite uniforme de fonctions simples. On remarquera d'autre part que $\mathsf{E}_{s,x} g(x_s(t,\ \omega))$ est une fonction $\mathfrak{B}$-mesurable en x pour toutes les g $\mathfrak{B}$-mesurables non négatives. Ceci découle de la mesurabilité de $\mathsf{E}_{s,x}\chi_B(x_s,(t,\ \omega))$ et partant de la mesurabilité de l'expression indiquée pour les g simples et les g limites de fonctions simples.

Soient $s < u < t_1 < \ldots < t_n$, $g_1, g_2, \ldots, g_n$ des fonctions $\mathfrak{B}$-mesurables bornées. On a

$$\mathsf{E}_{s,x}\Big(\prod_{k=1}^{n} g_k(x_s(t_k,\ \omega)) \mid \mathfrak{S}_u^s\Big) =$$

$$= \mathsf{E}_{u,\, x_s(u,\, \omega)} \prod_{k=1}^{n} g_k(x_u(t_k,\ \omega)) \pmod{\mathsf{P}_{s,x}}. \tag{6}$$

En effet, en utilisant les espérances mathématiques conditionnelles itérées on obtient

$$\mathsf{E}_{s,\, x}\Big(\prod_{k=1}^{n} g_k(x_s(t_k,\ \omega)) \mid \mathfrak{S}_u^s\Big) =$$

$$= \mathsf{E}_{s,\, x}\Big(\mathsf{E}_{s,\, x}\Big(\prod_{k=1}^{n} g_k(x_s(t_k,\ \omega)) \mid \mathfrak{S}_{t_{n-1}}^s\Big) \mid \mathfrak{S}_u^s\Big) =$$

$$= \mathsf{E}_{s,\, x}\Big(\prod_{k=1}^{n-1} g_s(x_u(t_k,\ \omega))\, \mathsf{E}_{t_{n-1},\, x_s(t_{n-1},\, \omega)} g_n(x_{t_{n-1}}(t_n,\ \omega)) \mid \mathfrak{S}_u^s\Big) .$$

Sous le signe de l'espérance mathématique on a maintenant un produit de $n-1$ fonctions si l'on admet la dernière égale à

$$g_{n-1}(x)\, \mathsf{E}_{t_{n-1},\, x} g_n(x_{t_{n-1}}(t_n,\ \omega)).$$

Donc en raisonnant par récurrence on trouve

$$\mathsf{E}_{s,x}\Big(\prod_{k=1}^{n} g_k(x_s(t,\omega))\,|\,\mathfrak{S}_u^s\Big)=$$

$$=\mathsf{E}_{s,x}(g_1(x_s(t_1,\omega))\,\mathsf{E}_{t_1,x_s(t_1,\omega)}g_2(x_{t_1}(t_2,\omega))\times$$

$$\times\,\mathsf{E}_{t_2,x_{t_1}(t_2,\omega)}g_3(x_{t_2}(t_3,\omega))\ldots\mathsf{E}_{t_{n-1},x_{t_{n-2}}(t_{n-1},\omega)}g_n(x_{t_{n-1}}(t_n,\omega))\,|\,\mathfrak{S}_u^s)=$$

$$=\mathsf{E}_{u,x_s(u,\omega)}g_1(x_u(t_1,\omega))\,\mathsf{E}_{t_1,x_u(t_1,\omega)}g_2(x_{t_1}(t_2,\omega))\times\ldots$$

$$\ldots\times\mathsf{E}_{t_{n-1},x_{t_{n-2}}(t_{n-1},\omega)}g_n(x_{t_{n-1}}(t_n,\omega))\pmod{\mathsf{P}_{s,x}}.\qquad(7)$$

En faisant $s=u$ on obtient

$$\mathsf{E}_{u,x}\Big(\prod_{k=1}^{n} g_k(x_u(t,\omega))\Big)=\mathsf{E}_{u,x}g_1(x_u(t,\omega))\times$$

$$\times\,\mathsf{E}_{t_1,x_u(t_1,\omega)}g_2(x_{t_1}(t_2,\omega))\ldots\mathsf{E}_{t_{n-1},x_{t_{n-2}}(t_{n-1},\omega)}g(x_{t_{n-1}}(t_n,\omega)).\qquad(8)$$

Si dans (8) on substitue $x_s(u,\omega)$ à x, on trouve l'expression du second membre de (6). Comme elle est confondue avec celle de (7), la formule (6) est démontrée.

Soient $\mathfrak{N}_t^s$ les tribus d'événements engendrées par $x_u(t',\omega)$, $u\leqslant s$, $t'\in[s,t]$. $\mathfrak{N}_t^s$ est une tribu d'événements engendrée par un observable $[s,t]$. On a $\mathfrak{N}_t^s\subset\mathfrak{S}_t^s$, puisque $x_u(t',\omega)$ est $\mathfrak{S}_t^s$-mesurable pour $u\leqslant s\leqslant t'\leqslant t$ d'après la condition e 2). Les tribus $\mathfrak{N}_t^s$ satisfont toutes les conditions imposées aux tribus $\mathfrak{S}_t^s$ de la définition du processus markovien (conditions b), e 2) et e 4)). Seule la condition e 4) est à vérifier. La formule des espérances mathématiques conditionnelles itérées donne

$$\mathsf{E}_{s,x}(g(x_s(t,\omega))\,|\,\mathfrak{N}_u^s)=\mathsf{E}_{s,x}(\mathsf{E}_{s,x}(g(x_s(t,\omega))\,|\,\mathfrak{S}_u^s\,|\,\mathfrak{N}_u^s)=$$

$$=\mathsf{E}_{s,x}(\mathsf{E}_{u,x_s(u,\omega)}g(x_u(t,\omega))\,|\,\mathfrak{N}_u^s)=\mathsf{E}_{u,x_s(u,\omega)}g(x_u(t,\omega))$$

$$\pmod{\mathsf{P}_{s,x}},\quad s\leqslant u\leqslant t,$$

puisque $\mathfrak{N}_u^s\subset\mathfrak{S}_u^s$ et $x_s(u,\omega)$ est $\mathfrak{N}_u^s$-mesurable et la fonction $\mathsf{E}_{u,x}g(x_u(t,\omega))$ est $\mathfrak{B}$-mesurable. On a prouvé que (5) a lieu si l'on remplace $\mathfrak{S}_u^s$ par $\mathfrak{N}_u^s$.

Donc si $\{x_t(s,\omega),\ \mathfrak{S}_t^s,\ \mathsf{P}_{s,x}\}$ est un processus markovien, alors $\{x_t(s,\omega),\ \mathfrak{N}_t^s,\ \hat{\mathsf{P}}_{s,x}\}$, où $\hat{\mathsf{P}}_{s,x}$ est la restriction de $\mathsf{P}_{s,x}$ à $\mathfrak{N}^s=\bigcup_t\mathfrak{N}_t^s$, l'est aussi. De toute évidence $\mathfrak{N}_t^s$ est la famille des plus petites tribus par rapport auxquelles $x_t(s,\omega)$ est markovien.

Voici enfin une relation généralisant (6) et exprimant la propriété d'être markovien dans sa forme la plus générale. Soit $\xi(\omega)$ variable aléatoire $\mathfrak{N}^t$-mesurable bornée. Pour $s\leqslant u\leqslant t$

$$\mathsf{E}_{s,x}(\xi(\omega)\,|\,\mathfrak{S}_u^s)=\mathsf{E}_{u,x_s(u,\omega)}\xi(\omega)\pmod{\mathsf{P}_{s,x}}.\qquad(9)$$

(9) résulte de (6) puisque toute variable $\mathfrak{N}^t$-mesurable bornée est limite stochastique (par rapport à $\mathsf{P}_{s,\,x}$) de sommes de la forme

$$\sum_m \prod_{k=1}^{n_m} g_k^{(m)}(x_s(t_k^{(m)},\ \omega)).$$

Soit, d'autre part, $\eta(\omega)$ variable $\mathfrak{S}_t^s$-mesurable. De (9) il résulte $\mathsf{E}_{s,\,x}\xi(\omega)\,\mathsf{E}_{s,\,x}(\eta(\omega)\,|\,\mathfrak{N}^t) = \mathsf{E}_{s,\,x}\eta(\omega)\,\xi(\omega) =$

$$= \mathsf{E}_{s,\,x}\eta(\omega)\,\mathsf{E}_{t,\,x_s(t,\,\omega)}\xi(\omega) = \mathsf{E}_{s,\,x}(\mathsf{E}_{t,\,x_s(t,\,\omega)}\xi(\omega))\,\mathsf{E}_{s,\,x}(\eta_\omega\,|\,\mathfrak{N}_t^t) =$$
$$= \mathsf{E}_{s,\,x}\xi(\omega)\,\mathsf{E}_{s,\,x}(\eta(\omega)\,|\,\mathfrak{N}_t^t),$$

donc

$$\mathsf{E}_{s,\,x}(\eta(\omega)\,|\,\mathfrak{N}^t) = \mathsf{E}_{s,\,x}(\eta(\omega)\,|\,\mathfrak{N}_t^t).$$

En vertu de (9), $\mathsf{E}_{s,\,x}(\xi(\omega)\,|\,\mathfrak{N}_t^s) = \mathsf{E}_{s,\,x}(\xi(\omega)\,|\,\mathfrak{N}_t^t)$. Par suite $\forall g(x)$ fonction mesurable bornée, on a

$\mathsf{E}_{s,\,x}\eta(\omega)\,g(x_s(t,\ \omega))\,\xi(\omega) =$

$$= \mathsf{E}_{s,\,x}\eta(\omega)\,g_s(x_s(t,\ \omega))\,\mathsf{E}_{s,\,x}(\xi(\omega)\,|\,\mathfrak{N}_t^t) =$$
$$= \mathsf{E}_{s,\,x}g_s(x_s(t,\ \omega))\,\mathsf{E}_{s,\,x}(\xi(\omega)\,|\,\mathfrak{N}_t^t)\,\mathsf{E}_{s,\,x}(\eta(\omega)\,|\,\mathfrak{N}_t^t).$$

D'où suit l'égalité

$$\mathsf{E}_{s,\,x}(\eta(\omega)\,\xi(\omega)\,|\,\mathfrak{N}_t^t) = \mathsf{E}_{s,\,x}(\eta(\omega)\,|\,\mathfrak{N}_t^t)\,\mathsf{E}_{s,\,x}(\xi(\omega)\,|\,\mathfrak{N}_t^t) \tag{10}$$

qui exprime l'indépendance du passé par rapport au futur, le présent étant donné.

Nous allons indiquer maintenant quelques extensions utiles des tribus $\mathfrak{N}_t^s$ conservant la propriété d'être markovien. Appelons $\mathfrak{N}_t^{s,\,x}$ la complétion de la tribu $\mathfrak{N}_t^s$ par rapport à la mesure $\mathsf{P}_{s,\,x}$, $\overline{\mathsf{P}}_{s,\,x}$ la complétion de la mesure,

$$\overline{\mathfrak{N}}_t^s = \bigcap_x \mathfrak{N}_t^{s,\,x}.$$

Alors $\{x_t(s,\ \omega),\ \overline{\mathfrak{N}}_t^s,\ \overline{\mathsf{P}}_{s,\,x}\}$ (on notera la restriction de la mesure à la plus petite tribu par le même symbole si aucune confusion n'est à craindre) est aussi processus markovien. Pour le prouver il suffit de vérifier que (5) est réalisée si l'on substitue $\overline{\mathfrak{N}}_u^s$ à $\mathfrak{S}_u^s$. Soit $\xi(\omega)$ fonction $\overline{\mathfrak{N}}_u^s$-mesurable bornée. Elle est $\mathfrak{N}_u^{s,\,x}$-mesurable. Donc existe $\xi_1(\omega)$ mesurable par rapport à $\mathfrak{N}_u^s$ telle que

$$\overline{\mathsf{P}}_{s,\,x}\{\xi(\omega) = \xi_1(\omega)\} = 1.$$

Par suite, pour $s < u < t$ ($\overline{\mathsf{E}}$ désigne l'intégrale par rapport à $\overline{\mathsf{P}}$), on aura

$\overline{\mathsf{E}}_{s,\,x}g(x_s(t,\ \omega))\,\xi(\omega) = \mathsf{E}_{s,\,x}g(x_s(t,\ \omega))\,\xi_1(\omega) =$

$$= \mathsf{E}_{s,\,x}\xi_1(\omega)\,\mathsf{E}_{s,\,x}(g(x_s(t,\ \omega))\,|\,\mathfrak{S}_u^s) =$$

$$= \mathsf{E}_{s,\,x}\xi_1(\omega)\,\mathsf{E}_{u,\,x_s(u,\,\omega)}g(x_u(t,\,\omega)) =$$
$$= \overline{\mathsf{E}}_{s,\,x}\xi(\omega)\,\mathsf{E}_{u,\,x_s(u,\,\omega)}g(x_u(t,\,\omega)).$$

D'où résulte (5).

Soient maintenant X espace métrique, $\mathfrak{B}$ tribu contenant toutes les boules. Posons

$$\mathfrak{N}^s_{t+} = \bigcap_{u>t}\mathfrak{N}^s_u.$$

THÉORÈME 1. *Supposons que la tribu $\mathfrak{B}$ est engendrée par un ensemble de fonctions continues et que la probabilité de passage $\mathsf{P}(s,\,x,\,t,\,B)$ d'un processus markovien $\{x_s(t,\,\omega),\,\mathfrak{N}^s_t,\,\mathsf{P}_{s,\,x}\}$ continu à droite est telle que: quels que soient s, t et $g(x)$ fonction $\mathfrak{B}$-mesurable continue bornée, la fonction*

$$g_{s,}(x) = \mathsf{E}_{s,\,x}g(x_s(t,\,\omega))$$

est continue en x et telle que

$$\lim_{\substack{h\downarrow 0\\ y\to x}}|g_{s+h}(x)-g_{s+h}(y)| = 0.$$

Le processus $\{x_s(t,\,\omega),\,\mathfrak{N}^s_{t+},\,\mathsf{P}_{s,\,x}\}$ est aussi markovien.

Démonstration. Il suffit de vérifier (5) si $\mathfrak{S}^s_t = \mathfrak{N}^s_{t+}$ pour les fonctions continues $g(x)$. Soit $\xi(\omega)$ fonction $\mathfrak{N}^s_{t+}$-mesurable bornée. Pour tout $h > 0$ elle est $\mathfrak{N}^s_{u+h}$-mesurable. Donc

$$\mathsf{E}_{s,\,x}\xi(\omega)\,g(x_s(t,\,\omega)) = \mathsf{E}_{s,\,x}\xi(\omega)\,\mathsf{E}_{s,\,x}(g(x_s(t,\,\omega))\,|\,\mathfrak{N}^s_{u+h}) =$$
$$= \mathsf{E}_{s,\,x}\xi(\omega)\,\mathsf{E}_{u+h,\,x_s(u+h,\,\omega)}g(x_{u+h}(t,\,\omega)) =$$
$$= \mathsf{E}_{s,\,x}\xi(\omega)\,g_{u+h}(x_s(u+h,\,\varphi)). \quad (11)$$

On remarquera que la fonction $g_u(x_s(u,\,\omega))$ est une martingale bornée, car pour $u_1 < u$

$$\mathsf{E}_{s,\,x}[g_u(x_s(u,\,\omega))\,\mathfrak{N}^s_{u_1}] = \mathsf{E}_{s,\,x}[[\mathsf{E}_{s,\,x}g(x_s(t,\,\omega))\,|\,\mathfrak{N}^s_u]/\mathfrak{N}^s_{u_1}] =$$
$$= \mathsf{E}_{s,\,x}[g(x_s(t,\,\omega))\,\mathfrak{N}^s_{u_1}] = g_{u_1}(x_s(u_1,\,\omega)).$$

Par suite, existe presque sûrement (cf. § 6, chapitre IV)

$$\lim_{h\downarrow 0} g_{u+h}(x_s(u+h,\,\omega)).$$

D'autre part,

$$\lim_{h\downarrow 0} g_{u+h}(x_s(u+h,\,\omega)) = \lim_{h\downarrow 0} g_{u+h}(x_s(u,\,\omega)),$$

puisque

$$\lim_{h\downarrow 0}[g_{u+h}(x_s(u+h,\,\omega)) - g_{u+h}(x_s(u,\,\omega))] = 0,$$

par hypothèse et d'après la continuité presque sûre à droite du processus. Soit η la valeur commune de ces limites. η est confondue

avec une variable $\mathfrak{N}_u^u$-mesurable comme limite de telles variables et par suite est de la forme $\lambda(x_s(u, \omega))$, où $\lambda(y)$ est une fonction $\mathfrak{B}$-mesurable. Dans (11), en passant à la limite pour $h \downarrow 0$, on trouve

$$\mathsf{E}_{s,x}\xi(\omega)\,g(x_s(t, \omega)) = \mathsf{E}_{s,x}\xi(\omega)\,\lambda(x_s(u, \omega)). \quad (12)$$

Si $\xi(\omega)$ est $\mathfrak{N}_u^s$-mesurable, le second membre de (12) est confondu avec l'expression

$$\mathsf{E}_{s,x}\xi(\omega)\,\mathsf{E}_{u,\,x_s(u,\,\omega)}g(x_s(t, \omega)).$$

Donc

$$\mathsf{E}_{s,x}\xi(\omega)\,\mathsf{E}_{u,\,x_s(u,\,\omega)}g(x_s(t, \omega)) = \mathsf{E}_{s,x}\xi(\omega)\,\lambda(x_s(u, \omega))$$

pour toutes les variables $\xi(\omega)$ $\mathfrak{N}_u^s$-mesurables bornées et par suite

$$\lambda(x_s(u, \omega)) = \mathsf{E}_{u,\,x_s(u,\,\omega)}g(x_s(t, \omega)).$$

En portant l'expression de λ dans (12) on s'assure que (5) et vraie pour $\mathfrak{S}_t^s = \mathfrak{N}_{t+}^s$. ■

Remarque. En utilisant la propriété d'être markovien du processus $\{x_s(t, \omega), \mathfrak{N}_{t+}^s, \mathsf{P}_{s,x}\}$, on établit une égalité analogue à (10) si $\eta(\omega)$ est $\mathfrak{N}_{t+}^s$-mesurable. Soit $B \in \mathfrak{N}_{t+}^t$. χ_B est $\mathfrak{N}_{t+}^s$-mesurable et $\mathfrak{N}^t$-mesurable. Par suite

$$\mathsf{P}_{s,x}(B \mid \mathfrak{N}_t^t) = \mathsf{E}_{s,x}(\chi_B\chi_B \mid \mathfrak{N}_t^t) =$$
$$= \mathsf{E}_{s,x}(\chi_B \mid \mathfrak{N}_t^t)\,\mathsf{E}_{s,x}(\chi_B \mid \mathfrak{N}_t^t) = \mathsf{P}_{s,x}^2(B \mid \mathfrak{N}_t^t).$$

Donc, *dans les hypothèses du théorème* 1, $\mathsf{P}_{s,x}(B \mid \mathfrak{N}_t^t)$ *est égale ou à* 0 *ou à* 1 *pour* $B \in \mathfrak{N}_{t+}^t$. *Cette assertion porte le nom de loi de tout ou rien pour les processus markoviens.*

On utilisera dans la suite la notion d'équivalence stochastique de processus markoviens. Soient donnés deux processus markoviens $\{x_s(t, \omega), \mathfrak{S}_t^s, \mathsf{P}_{s,x}\}$ et $\{\bar{x}_s(t, \omega), \bar{\mathfrak{S}}_t^s, \bar{\mathsf{P}}_{s,x}\}$ ayant le même espace des phases $(X, \mathfrak{B})$ et le même espace d'événements élémentaires $\{\Omega, \mathfrak{S}\}$. On dit qu'ils sont stochastiquement équivalents si existe un processus markovien $\{\tilde{x}_s(t, \omega), \tilde{\mathfrak{S}}_t^s, \tilde{\mathsf{P}}_{s,x}\}$ tel que

a) $\tilde{\mathfrak{S}}_t^s \subset \mathfrak{S}_t^s$, $\tilde{\mathfrak{S}}_t^s \subset \bar{\mathfrak{S}}_t^s$;

b) les mesures $\tilde{\mathsf{P}}_{s,x}$, $\bar{\mathsf{P}}_{s,x}$ et $\mathsf{P}_{s,x}$ sont confondues sur $\tilde{\mathfrak{S}}^s$;

c) $\mathsf{P}_{s,x}\{x_s(t, \omega) = \tilde{x}_s(t, \omega)\} = 1$, $\bar{\mathsf{P}}_{s,x}\{\bar{x}_s(t, \omega) = \tilde{x}_s(t, \omega)\} = 1$

pour tous les t. Les processus stochastiquement équivalents possèdent visiblement les mêmes probabilités de passage.

Soit $\{x_s(t, \omega), \mathfrak{S}_t^s, \mathsf{P}_{s,x}\}$ processus markovien. Une variable aléatoire non négative τ est par définition *instant markovien* pour

ce processus si l'événement $\{\tau > t\} \in \mathfrak{S}_t^0$ pour $t > 0$, i.e. si en observant le processus sur l'intervalle $[0, t]$ on peut dire si l'instant τ a eu lieu ou non. D'une façon générale, on envisagera des instants susceptibles de prendre la valeur $+\infty$. Outre les instants markoviens on étudiera des instants s-markoviens ($s \geqslant 0$) ; dans ce cas $\tau \geqslant s$ et l'événement $\{\tau > t\} \in \mathfrak{S}_t^s$ pour $t \geqslant s$. Les instants 0-markoviens sont simplement markoviens. Soit τ instant s-markovien. Appelons $\mathfrak{S}_\tau^s$ l'ensemble des événements de $\mathfrak{S}^s$ tels que

$$A \cap \{\tau \leqslant t\} \in \mathfrak{S}_t^s.$$

De toute évidence, $\mathfrak{S}_\tau^s$ est une tribu. τ est visiblement mesurable par rapport à $\mathfrak{S}_\tau^s$.

LEMME 1. *Si X est un espace métrique, $\mathfrak{B}$ une tribu engendrée par une classe de fonctions continues et $x_s(t, \omega)$ continue à droite, alors $x_s(\tau, \omega)$ est aussi $\mathfrak{S}_\tau^s$-mesurable pour tout instant τ s-markovien.*

Démonstration. Pour toute fonction $g(x)$ mesurable continue on a

$$g(x_s(\tau, \omega))\, \chi_{\{\tau \leqslant t\}} = \lim_{n \to \infty} \sum_{k=1}^{2^n} g\left(x_s\left(t\frac{k}{2^n}, \omega\right)\right) \chi_{\left\{\frac{k-1}{2^n} t < \tau \leqslant \frac{k}{2^n} t\right\}}.$$

Donc $g(x_s(\tau, \omega))\, \chi_{\{\tau \leqslant t\}}$ est $\mathfrak{S}_t^s$-mesurable. Par suite, ceci est valable pour toute fonction mesurable g et en particulier pour les B $\mathfrak{B}$-mesurables :

$$\chi_B(x_s(\tau, \omega))\, \chi_{\{\tau \leqslant t\}}$$

est $\mathfrak{S}_t^s$-mesurable et $\{x_s(\tau, \omega) \in B\} \cap \{\tau \leqslant t\} \in \mathfrak{S}_t^s$, $\{x_s(\tau, \omega) \in B\} \in \mathfrak{S}_\tau^s$. ∎

Les exemples les plus intéressants d'instants markoviens sont les instants de première atteinte (resp. de sortie) d'un ensemble.

Soient X espace métrique et $x_s(t, \omega)$ continue à droite, G ensemble ouvert. Alors la quantité $\tau = \sup[t : x_s(u, \omega) \notin G,\ u \leqslant t]$ appelée *instant de première atteinte de G* sera instant s-markovien. En effet,

$$\{\tau > t\} = \bigcap_k \{\omega : x_s(t_k, \omega) \notin G\} \cap \{\omega : x_s(t, \omega) \notin G\},$$

où $\{t_k\}$ est un ensemble partout dense dans $[s, t]$ et les ensembles figurant dans le second membre appartiennent à $\mathfrak{S}_t^s$. Si $\tau_n \uparrow \tau$ et τ_n sont des instants s-markoviens, alors τ le sera également :

$$\{\tau > t\} = \bigcup_n \{\tau_n > t\}.$$

Donc pour tout ensemble $F = \bigcap_n G_n$, où G_n est une suite monotone décroissante d'ouverts,

$$\tau = \sup[t : x_s(u, \omega) \notin F,\ u \leqslant t]$$

sera aussi s-markovien, puisque

$$\tau = \sup_n \tau_n,$$

où τ_n est l'instant de première atteinte de l'ensemble G_n. En particulier, l'instant de première atteinte d'un fermé est aussi s-markovien.

Les processus markoviens qui le restent aux instants markoviens forment une classe très importante. On les appelle *processus markoviens forts*. Donnons la définition exacte de cette notion.

DÉFINITION. *Un processus* $\{x_s(t, \omega), \mathfrak{S}_t^s, \mathsf{P}_{s,x}\}$ *est markovien fort si*

I. $x_s(\tau, \omega)$ *est* $\mathfrak{S}_\tau^s$*-mesurable pour tout instant* τ *s-markovien.*

II.

$$\mathsf{E}_{s,x}(g(x_s(\tau+t, \omega)) \mid \mathfrak{S}_\tau^s) = \\ = \mathsf{E}_{\tau, x_s(\tau, \omega)} g(x_\tau(\tau+t, \omega)) \pmod{\mathsf{P}_{s,x}}, \quad (13)$$

$g(x)$ étant $\mathfrak{B}$-mesurable, $t > 0$.

Le second membre de (13) est le résultat de la substitution $u = \tau$, $x = x_s(\tau, \omega)$ dans la fonction $\mathsf{E}_{u,x} g(x_u(u+t, \omega))$. Pour que le résultat de cette substitution soit variable aléatoire il faut que:

III. *Pour toute fonction mesurable bornée g la fonction* $\mathsf{E}_{u,x} g(x_u(u+t, \omega))$ *soit mesurable en l'ensemble des variables u, x par rapport au produit des tribus* $\mathfrak{A} \times \mathfrak{B}$, *où* $\mathfrak{A}$ *est la tribu des boréliens de* $[0, \infty[$.

LEMME 2. *La relation* (13) *est réalisée pour tout instant s-markovien prenant un nombre dénombrable de valeurs.*

Démonstration. Soient $t_1, t_2, \ldots, t_n, \ldots$ l'ensemble de toutes les valeurs prises par τ. On a $\{\tau = t_k\} \in \mathfrak{S}_{t_k}^s$. Soit $A \in \mathfrak{S}_\tau^s$. Il vient $A \cap \{\tau = t_k\} \in \mathfrak{S}_{t_k}^s$ et

$$\mathsf{E}_{s,x} \chi_A g(x_s(\tau+t, \omega)) = \sum_k \mathsf{E}_{s,x} \chi_A \chi_{\{\tau=t_k\}} g(x_s(\tau+t, \omega)) =$$

$$= \sum_k \mathsf{E}_{s,x} \chi_A \chi_{\{\tau=t_k\}} g(x_s(t_k+t, \omega)) =$$

$$= \sum_k \mathsf{E}_{s,x} \chi_{\{\tau=t_k\} \cap A} \mathsf{E}_{s,x}(g(x_s(t_k+t, \omega)) \mid \mathfrak{S}_{t_k}^s) =$$

$$= \sum_k \mathsf{E}_{s,x} \chi_{\{\tau=t_k\}} \chi_A \mathsf{E}_{t_k, x_s(t_k, \omega)} g(x_{t_k}(t_k+t, \omega)) =$$

$$= \sum_k \mathsf{E}_{s,\,x}\chi_A\chi_{\{\tau=t_k\}}\mathsf{E}_{\tau,\,x_s(\tau,\,\omega)}g\,(x_\tau\,(\tau+t,\ \omega)) =$$
$$= \mathsf{E}_{s,\,x}\chi_A\mathsf{E}_{\tau,\,x_s(\tau,\,\omega)}g\,(x_\tau\,(\tau+t,\ \omega)),$$

d'où découle (13). ■

Voici une autre condition suffisante d'être markovien fort.

THÉORÈME 2. *Soient X espace métrique,* $\mathfrak{B}$ *la tribu engendrée par une classe de fonctions continues et* $\{x_s\,(t,\ \omega),\ \mathfrak{S}_t^s,\ \mathsf{P}_{s,\,x}\}$ *processus markovien tel que:*

1) $x_s\,(t,\ \omega)$ *est continue à droite en* t;

2) *pour toute fonction* $f\,(x)$ *continue bornée mesurable, la fonction*

$$R_\lambda f\,(s,\ x) = \int_0^\infty e^{-\lambda t} \int \mathsf{P}\,(s,\ x,\ s+t,\ dy)\,f\,(y) \tag{14}$$

est telle que

$$\lim_{\substack{h\downarrow 0\\ y\to x}} R_\lambda f\,(s+h,\ y) = R_\lambda f\,(s,\ x).$$

Alors le processus $\{x_s\,(t,\ \omega),\ \mathfrak{S}_t^s,\ \mathsf{P}_{s,\,x}\}$ *est markovien fort.*

D é m o n s t r a t i o n. Soit τ instant s-markovien. Posons

$$\tau_n = \frac{k+1}{n}\ \text{si}\ \frac{k}{n} < \tau \leqslant \frac{k+1}{n}.$$

De toute évidence, $\tau_n \geqslant \tau$ et

$$\{\tau_n \leqslant t\} = \left\{\tau_n \leqslant \frac{[nt]}{n}\right\} = \left\{\tau \leqslant \frac{[nt]}{n}\right\} \in \mathfrak{S}^s_{\frac{[nt]}{n}} \subset \mathfrak{S}_t^s$$

($[x]$ est la partie entière de x). Donc τ_n est aussi s-markovien. Par suite, en vertu du lemme 2,

$$\mathsf{E}_{s,\,x}\,(g\,(x_s\,(\tau_n+t,\ \omega))\mid \mathfrak{S}^s_{\tau_n}) =$$
$$= \mathsf{E}_{\tau_n,\,x_s(\tau_n,\,\omega)}g\,(x_{\tau_n}\,(\tau_n+t,\ \omega))\ (\operatorname{mod} \mathsf{P}_{s,\,x}). \tag{15}$$

Le premier membre de (15) est confondu avec

$$\mathsf{E}_{u,\,x}g\,(x_u\,(u+t,\ \omega)) \tag{16}$$

si l'on pose $u = \tau_n$, $x = x_s\,(\tau_n,\ \omega)$. L'expression (16) est continue à droite et par conséquent intégrable, puisque $x_u\,(u+t,\ \omega)$ est continue à droite en t pour g continues. Donc les fonctions de (15)

sont intégrables par rapport à t et par suite

$$\int_0^\infty e^{-\lambda t} \mathsf{E}_{s,x}(g(x_s(\tau_n+t,\ \omega)) \mid \mathfrak{S}^s_{\tau n})\,dt =$$

$$= \int_0^\infty e^{-\lambda t} \mathsf{E}_{\tau n,\ x_s(\tau n,\ \omega)} g(x_{\tau n}(\tau_n+t,\ \omega))\,dt) \pmod{\mathsf{P}_{s,x}}. \qquad (17)$$

Soit $A \in \mathfrak{S}^s_\tau$. On a

$$A \cap \{\tau_n \leqslant t\} = A \cap \left\{\tau_n \leqslant \frac{[nt]}{n}\right\} = A \cap \left\{\tau \leqslant \frac{[nt]}{n}\right\} \in \mathfrak{S}^s_t.$$

Donc $A \in \mathfrak{S}^s_{\tau n}$. En multipliant (17) par χ_A et en prenant $\mathsf{E}_{s,x}$, on trouve

$$\int_0^\infty e^{-\lambda t} \mathsf{E}_{s,x} g(x_s(\tau_n+t,\ \omega))\,\chi_A\,dt = \mathsf{E}_{s,x} R_\lambda g(\tau_n,\ x_s(\tau_n,\ \omega))\,\chi_A.$$

En passant à la limite pour $n \to \infty$ et eu égard à $\tau_n \geqslant \tau$, $\tau_n \to \tau$, $x_s(\tau_n + t,\ \omega) \to x_s(\tau + t,\ \omega)$ et à la condition 2) du théorème, on trouve

$$\int_0^\infty e^{-\lambda t} \mathsf{E}_{s,x} g(x_s(\tau+t,\ \omega))\,\chi_A\,dt = \mathsf{E}_{s,x} R_\lambda g(\tau,\ x_s(\tau,\ \omega))\,\chi_A =$$

$$= \mathsf{E}_{s,x} \int_0^\infty e^{-\lambda t} \mathsf{E}_{\tau,\ x_s(\tau,\ \omega)} g(x_\tau(\tau+t,\ \omega))\,\chi_A\,dt =$$

$$= \int_0^\infty e^{-\lambda t} \mathsf{E}_{s,x} \mathsf{E}_{\tau,\ x_s(\tau,\ \omega)} g(x_\tau(\tau+t,\ \omega))\,\chi_A\,dt.$$

On sait que deux fonctions continues à droite dont les transformées de Laplace sont identiques sont confondues. Donc

$$\mathsf{E}_{s,x} g(x_s(\tau+t,\ \omega))\,\chi_A = \mathsf{E}_{s,x} \mathsf{E}_{\tau,\ x_s(\tau,\ \omega)} g(x_\tau(\tau+t,\ \omega))\,\chi_A \qquad (18)$$

pour tout $A \in \mathfrak{S}^s_\tau$. Par suite, (13) est réalisée pour toutes les fonctions g mesurables continues bornées et par voie de conséquence pour toutes les fonctions g mesurables bornées. La $\mathfrak{S}^s_\tau$-mesurabilité de $x_s(\tau,\ \omega)$ résulte du lemme 1. ■

La relation (13) se généralise de la façon suivante pour les processus markoviens forts.

Soient $t_1 < t_2$, $g_1(x)$ et $g_2(x)$ deux fonctions mesurables bornées, τ instant s-markovien. Alors $\tau + t_1$ est aussi s-markovien: pour $u > t_1 + s$

$$\{\tau + t_1 \leqslant u\} = \{\tau \leqslant u - t_1\} \in \mathfrak{S}^s_{u-t_1} \subset \mathfrak{S}^s_u.$$

On s'assure immédiatement que $\mathfrak{S}^s_{\tau+t_1} \supset \mathfrak{S}^s_\tau$. Ecrivons la relation (13) pour la fonction g_2 et l'instant $\tau+t_1$:

$$\mathsf{E}_{s,x}(g_2(x_s(\tau+t_2,\omega))|\mathfrak{S}^s_{\tau+t_1}) = \mathsf{E}_{\tau+t_1,\,x_s(\tau+t_1,\,\omega)}g_2(x_{\tau+t_1}(\tau+t_2,\omega)).$$

En multipliant cette égalité par $g_1(x_s(\tau+t_1,\omega))$ et en prenant l'espérance mathématique conditionnelle par rapport à $\mathfrak{S}^s_\tau$, on trouve

$$\mathsf{E}_{s,x}(g_1(x_s(\tau+t_1,\omega))\,g_2(x_s(\tau+t_2,\omega))\mid\mathfrak{S}^s_\tau) =$$
$$= \mathsf{E}_{s,x}(g_1(x_s(\tau+t_1,\omega))\,\mathsf{E}_{\tau+t_1,\,x_s(\tau+t_1,\,\omega)}g_2(x_{\tau+t_1}(\tau+t_2,\omega))\mid\mathfrak{S}^s_\tau).$$

Utilisons maintenant l'égalité suivante: si $f(x, s)$ est fonction $\mathfrak{B}\times\mathfrak{A}$-mesurable bornée, τ instant s-markovien, alors

$$\mathsf{E}_{s,x}(f(x_s(\tau+t,\omega),\tau)\mid\mathfrak{S}^s_\tau) =$$
$$= \mathsf{E}_{\tau,\,x_s(\tau,\,\omega)}f(x_\tau(\tau+t,\omega),\tau) \quad (\mathrm{mod}\ \mathsf{P}_{s,x}) \qquad (18')$$

(le second membre de (18′) est le résultat de la substitution $u = \tau$, $x = x_s(\tau,\omega)$ dans la fonction $\mathsf{E}_{u,\,x}f(x_u(u+t,\omega),u)$). L'égalité (18′) est évidente, si $f(x, u) = f_1(x)\,g_1(u)$, en vertu de la $\mathfrak{S}^s_\tau$-mesurabilité de τ et de (13). On peut utiliser maintenant le fait que la fonction $f(x, u)$ est limite par rapport à une mesure quelconque d'une $\mathfrak{B}\times\mathfrak{A}$-suite de fonctions de la forme

$$\sum_k f_k(x)\,g_k(u).$$

En se servant de (18) on obtient

$$\mathsf{E}(g_1(x_s(\tau+t_1,\omega))\,\mathsf{E}_{\tau+t_1,\,x_s(\tau+t_1,\,\omega)}\,g_2(x_{\tau+t_1}(\tau+t_2,\omega))\mid\mathfrak{S}^s_\tau) =$$
$$= \mathsf{E}_{\tau,\,x_s(\tau,\,\omega)}g_1(x_\tau(\tau+t_1,\omega))\,\mathsf{E}_{\tau+t_1,\,x_\tau(\tau+t_1,\,\omega)}g_2(x_{\tau+t_1}(\tau+t_2,\omega)).$$

Donc

$$\mathsf{E}_{s,x}(g_1(x_s(\tau+t_1,\omega))\,g_2(x_s(\tau+t_2,\omega))\mid\mathfrak{S}^s_\tau) =$$
$$= \mathsf{E}_{\tau,\,x_s(\tau,\,\omega)}g_1(x_\tau(\tau+t_1,\omega))\times$$
$$\times\,\mathsf{E}_{\tau+t_1,\,x_\tau(\tau+t_1,\,\omega)}g_2(x_{\tau+t_1}(\tau+t_2,\omega)) \quad (\mathrm{mod}\ \mathsf{P}_{s,x}).$$

On établit de façon analogue la formule suivante: pour $0 < t_1 < t_2 < \ldots < t_k$, $g_1(x), \ldots, g_k(x)$, fonctions mesurables bornées, et τ, instant s-markovien, on a

$$\mathsf{E}_{s,x}\Big(\prod_{j=1}^{k} g_j(x_s(\tau+t_j,\omega))\mid\mathfrak{S}^s_\tau\Big) =$$
$$= \mathsf{E}_{\tau,\,x_s(\tau,\,\omega)}g_1(x_\tau(\tau+t_1,\omega))\,\mathsf{E}_{\tau+t_1,\,x_\tau(\tau+t_1,\,\omega)}g_2(x_{\tau+t_1}(\tau+t_2,\omega))\ldots$$
$$\ldots\mathsf{E}_{\tau+t_{k-1},\,x_{\tau+t_{k-2}}}(\tau+t_{k-1},\omega)\,g_k(x_{\tau+t_{k-1}}(\tau+t_k,\omega))$$
$$(\mathrm{mod}\ \mathsf{P}_{s,x}). \qquad (19)$$

De (19) et compte tenu de (8) et (6) on obtient la formule suivante:

$$\mathsf{E}_{s,x}\Big(\prod_{j=1}^{k} g_j(x_s(\tau+t_j,\omega))\mid \mathfrak{S}_\tau^s\Big)=$$

$$=\mathsf{E}_{\tau,\,x_s(\tau,\,\omega)}\prod_{j=1}^{k} g_j(x_\tau(\tau+t_j,\omega))\quad (\mathrm{mod}\ \mathsf{P}_{s,x}). \tag{20}$$

§ 2. Processus markoviens discontinus généraux

Soit $\{X, \mathfrak{B}\}$ espace probabilisable, $\mathfrak{B}$ est une tribu contenant tous les ensembles à un point. Un processus markovien $\{x_s(t,\omega), \mathfrak{S}_t^s, \mathsf{P}_{s,x}\}$ est *en escalier* si

$$\lim_{\delta\downarrow 0} \mathsf{P}_{s,x}(\{\omega : x_s(t,\omega)=x,\ s\leqslant t\leqslant s+\delta\})=1 \tag{1}$$

pour tous les s et x. Si la condition (1) est réunie, alors pour tous les s est définie une variable aléatoire positive

$$\tau_1=\sup\,[t:\ x_s(u,\omega)=x_s(s,\omega),\quad s\leqslant u\leqslant t]$$

appelée instant de première sortie de l'état initial. Cette quantité est visiblement instant s-markovien.

On remarquera que ω, l'ensemble figurant sous le signe de la probabilité dans (1), n'est pas forcé d'être événement, puisqu'il est intersection d'un continuum d'événements. Pour qu'il soit événement, on considérera uniquement les processus dont les réalisations sont séparables au sens suivant: si $x_s(t,\omega)=x$ pour $t\in\ ]\alpha,\beta[\ \cap\Lambda$, où Λ est un ensemble dénombrable partout dense dans $[0,\infty[$, alors $x_s(t,\omega)=x$ pour tous les $t\in[\alpha,\beta[$.

La condition (1) entraîne en particulier

$$\lim_{t\downarrow s} \mathsf{P}_{s,x}(\{\omega : x_s(t,\omega)=x\})=\lim_{t\downarrow s}\mathsf{P}(s,x,t,\{x\})=1. \tag{2}$$

Les processus dont les probabilités de passage vérifient la condition (2) $\forall s$, $\forall x$, sont dits *stochastiquement continus.*

LEMME 1. *Un processus stochastiquement continu est stochastiquement équivalent à un processus séparable.*

Démonstration. Construisons un processus $\tilde{x}_s(t,\omega)$ de la façon suivante:

$$\tilde{x}_s(t,\omega)=\begin{cases} x \text{ si } \exists\delta>0 \text{ tel que } x_s(u,\omega)=x \text{ pour } u\in\,]t,t+\delta[\,\cap\Lambda;\\ x_s(t,\omega) \text{ dans les autres cas.}\end{cases}$$

Appelons $\mathfrak{A}_k$ l'événement $\{x_s(t_i,\omega)=x_s(t_j,\omega)\}$, $t_i, t_j\in\,\left]t, t+\frac{1}{k}\right[$, $\Lambda=\{t_1,t_2,\ldots\}$. Choisissons une suite $t_{n_k}\in\Lambda$ telle que

$t < t_{n_k} < t + \frac{1}{k}$. Alors

$$\{\widetilde{x}_s(t, \omega) \neq x_s(t, \omega)\} \subset \bigcup_k \mathfrak{A}_k.$$

Donc

$$\mathsf{P}_{s,x}(\{\widetilde{x}_s(t, \omega) \neq x_s(t, \omega)\}) \leqslant \lim_{k\to\infty} \mathsf{P}_{s,x}(\{\widetilde{x}_s(t, \omega) \neq$$

$$\neq x_s(t, \omega)\} \cap \mathfrak{A}_k) \leqslant \lim_{k\to\infty} \mathsf{P}_{s,x}(\{x_s(t_{n_k}, \omega) \neq x_s(t, \omega)\} \cap \mathfrak{A}_k) \leqslant$$

$$\leqslant \lim_{k\to\infty} \mathsf{P}_{s,x}(x_s(t_{n_k}, \omega) \neq x_s(t, \omega)) =$$

$$= \lim_{k\to\infty} \mathsf{E}_{s,x}\mathsf{E}_{t, x_s(t,\omega)}\chi\{x_t(t, \omega) \neq x_t(t_{n_k}, \omega)\} =$$

$$= \lim_{k\to\infty} \int \mathsf{P}(s, x, t, dy)\,\mathsf{P}(t, y, t_{n_k}, X - \{y\}) = 0,$$

puisque $t_{n_k} \downarrow t$ et $\mathsf{P}(t, y, t_{n_k}, X - \{y\}) \to 0$ pour tous les y. De toute évidence, $\widetilde{x}_s(t, \omega)$ sera mesurable par rapport à la tribu des $\widetilde{\mathfrak{S}}_t^s$-événements, engendrée par les événements de $\mathfrak{S}_t^s$ et $\bigcap_m \widetilde{N}^t_{t+\frac{1}{m}}$, où $\widetilde{N}^t_{t+\frac{1}{m}}$ est engendrée par l'ensemble $\mathfrak{A}_m \cap A_{t_k}$, $A_{t_k} \subset \mathfrak{N}_{t_k}^{t_k}$, $t_k \in]t, t+\delta[$. Comme $\widetilde{\mathfrak{S}}_t^s \subset \mathfrak{S}_t^s$, la mesure $\mathsf{P}_{s,x}$ est définie sur $\widetilde{\mathfrak{S}}_t^s$. Montrons que $\{x_s(t, \omega), \widetilde{\mathfrak{S}}_t^s, \mathsf{P}_{s,x}\}$ est aussi processus markovien.

Soit ξ variable aléatoire $\widetilde{\mathfrak{S}}_t^s$-mesurable bornée. $\widetilde{\mathfrak{S}}_t^s$ étant contenue dans la complétion de $\mathfrak{S}_t^s$ par rapport à la mesure $\mathsf{P}_{s,x}$ il existe $\bar{\xi}$, variable aléatoire $\mathfrak{S}_t^s$-mesurable, telle que $\mathsf{P}_{s,x}\{\xi = \bar{\xi}\} = 1$. Par suite, pour $u > t$

$$\mathsf{E}_{s,x}\xi\chi_B(x_s(u, \omega)) = \mathsf{E}_{s,x}\bar{\xi}\chi_B(x_s(u, \omega)) =$$

$$= \mathsf{E}_{s,x}\bar{\xi}\mathsf{E}_{t, x_s(t,\omega)}\chi_B(x_t(u, \omega)) =$$

$$= \mathsf{E}_{s,k}\xi\mathsf{E}_{t, x_s(t,\omega)}\chi_B(x_t(u, \omega)).$$

D'où suit l'égalité

$$\mathsf{P}_{s,x}(\{x_s(u, \omega) \in B\} \mid \widetilde{\mathfrak{S}}_t^s) = \mathsf{P}_{t, x_s(t,\omega)}(\{x_t(u, \omega) \in B\})$$

Or, en utilisant l'égalité $\mathsf{P}_{s,x}\{x_s(u, \omega) = \widetilde{x}_s(u, \omega)\} = 1$, on s'assure que la dernière relation est réalisée si l'on remplace $x_s(\cdot, \cdot)$ par $\widetilde{x}_s(\cdot, \cdot)$. ■

On supposera dans la suite que tous les processus sont stochastiquement continus et séparables. Trouvons la condition de réalisation de (1).

THÉORÈME 1. *Pour qu'un processus markovien séparable d'espace des phases $(X, \mathfrak{B})$ soit en escalier, il faut et il suffit que pour tout $s \in [0, \infty[$, $x \in X$*

$$\lim_{\delta \downarrow 0} \overline{\lim_{\max(t_k^{(n)} - t_{k-1}^{(n)}) \to 0}} \sum_{s \leqslant t_0^{(n)} < \ldots < t_n^{(n)} = s+\delta} \mathsf{P}(t_k^{(n)}, x, t_{k+1}^{(n)}, X - \{x\}) = 0, \quad (3)$$

quelle que soit $s = t_0^{(n)} < \ldots < t_n^{(n)} = s + \delta$, suite de partitions telle que $\max(t_k^{(n)} - t_{k-1}^{(n)}) \to 0$.

D é m o n s t r a t i o n. *Condition nécessaire.* Soit réalisée la condition (1). On a

$$\mathsf{P}_{s,x}(\{\omega : x_s(t, \omega) = x, s \leqslant t \leqslant s + \delta\}) =$$

$$= \lim_{\max(t_k^{(n)} - t_{k-1}^{(n)}) \to 0} \mathsf{P}_{s,x}\{x_s(t_k^{(n)}, \omega) = x, k = \overline{1, n}\} =$$

$$= \lim_{\max(t_k^{(n)} - t_{k-1}^{(n)}) \to 0} \prod_{k=1}^{n} \mathsf{P}(t_{k-1}^{(n)}, x, t_k^{(n)}, \{x\}) =$$

$$= \lim_{\max(t_k^{(n)} - t_{k-1}^{(n)}) \to 0} \prod_{k=1}^{n} (1 - \mathsf{P}(t_{k-1}^{(n)}, x, t_k^{(n)}, X - \{x\})) \leqslant$$

$$\leqslant \exp\Big\{- \lim_{\max(t_k^{(n)} - t_{k-1}^{(n)}) \to 0} \sum_{k=1}^{n} \mathsf{P}(t_{k-1}^{(n)}, x, t_k^{(n)}, X - \{x\})\Big\}.$$

Donc

$$\overline{\lim_{\max(t_k^{(n)} - t_{k-1}^{(n)}) \to 0}} \sum_{k=1}^{n} \mathsf{P}(t_{k-1}^{(n)}, x, t_k^{(n)}, X - \{x\}) \leqslant$$

$$\leqslant - \ln \mathsf{P}_{s,x}(\{\omega : x_s(t, \omega) = x, s \leqslant t \leqslant s + \delta\}).$$

En passant à la limite pour $\delta \downarrow 0$, de (1) on déduit (3).

Condition suffisante. Soit réalisée (3). Choisissons une suite de partitions $s = t_0^{(n)} < t_1^{(n)} < \ldots < t_n^{(n)} = s + \delta$, telle que les ensembles $\Lambda_n = \{t_1^{(n)}, \ldots t_{n-1}^{(n)}\}$ soient monotones croissants avec n et

$\bigcup_n \Lambda_n = \Lambda \cap (s, s+\delta)$. Il vient

$$\mathsf{P}_{s,x}(\{x_s(t, \omega) = x, s \leqslant t \leqslant s+\delta\}) =$$
$$= \mathsf{P}_{s,x}(\{x_s(t, \omega) = x, t \in \Lambda \cap (s, s+\delta)\}) =$$
$$= \lim_{n\to\infty} \mathsf{P}_{s,x}(\{x_s(t, \omega) = x, t \in \Lambda_n\}) =$$
$$= \lim_{n\to\infty} \prod_{k=1}^{n-1} \mathsf{P}(t_{k-1}^{(n)}, x, t_k^{(n)}, \{x\}) =$$
$$= \lim_{n\to\infty} \prod_{k=1}^{n-1} (1 - \mathsf{P}(t_{k-1}^{(n)}, x, t_k^{(n)}, X - \{x\})) =$$
$$= \varliminf_{n\to\infty} \exp\left\{\sum_{k=1}^{n-1} \ln(1 - \mathsf{P}(t_{k-1}^{(n)}, x, t_k^{(n)}, X - \{x\}))\right\}. \quad (4)$$

Si δ est choisi tel que

$$\varliminf \sum_{k=1}^{n} \mathsf{P}(t_{k-1}^{(n)}, x, t_k^{(n)}, X - \{x\}) < \varepsilon,$$

alors

$$\varliminf \sup_k \mathsf{P}(t_{k-1}^{(n)}, x, t_k^{(n)}, X - \{x\}) < \varepsilon$$

et

$$\varliminf_{n\to\infty} \sum_{k=1}^{n-1} \ln(1 - \mathsf{P}(t_{k-1}^{(n)}, x, t_k^{(n)}, X - \{x\})) \geqslant$$
$$\geqslant [1 + O(\varepsilon)] \varliminf_{n\to\infty} \sum_{k=1}^{n} \mathsf{P}(t_{k-1}^{(n)}, x, t_k^{(n)}, X - \{x\}).$$

Donc, compte tenu de (3), on déduit (1) de (4). ■

Soit dans X une topologie discrète dans laquelle $\{x\}$ est le plus petit voisinage du point x. Toute fonction définie sur X sera alors continue. Les relations

$$\mathsf{P}(s, x, t+h, A) = \int \mathsf{P}(s, x, t, dy)\, \mathsf{P}(t, y, t+h, A)$$

et

$$\lim_{h\downarrow 0} \mathsf{P}(t, y, t+h, A) = \chi_A(y),$$

$$\mathsf{P}(s, x, s+h, \{x\})\,\mathsf{P}(s+h, x, t, A)+ \\ + \int\limits_{X-\{x\}} \mathsf{P}(s, x, s+h, dy)\,\mathsf{P}(s+h, y, t, A) = \mathsf{P}(s, x, t, A)$$

entraînent la continuité à droite en s et en t de $\mathsf{P}(s, x, t, A)$ pour tout processus stochastiquement continu. Donc la fonction

$$\int\limits_0^\infty e^{-\lambda t} \int \mathsf{P}(s, x, t+s, dy)\, g(y)\, dt$$

est continue à droite en s. Le processus étant en escalier, il s'ensuit qu'il prend chacune de ses valeurs sur un intervalle que l'on peut supposer être fermé à gauche (tel est le processus construit dans le lemme 1). Cependant ceci n'en garantit pas la continuité à droite comme le montre l'exemple

$$x(t) = k \text{ pour } \frac{1}{2^k} \leqslant t < \frac{1}{2^{k-1}}, \qquad x(0) = 0.$$

Dans une topologie discrète, les processus continus à droite seront nécessairement en escalier. Ils forment une classe plus étroite que les processus en escalier. Nous les nommerons *processus discontinus.* Le théorème 2, § 1, nous dit que tout processus discontinu est markovien fort.

Considérons les instants markoviens τ_k correspondant au saut de rang k du processus $x_s(t, \omega)$. Ils se définissent par récurrence de la manière suivante: τ_1 est l'instant de sortie de l'état initial, on l'a défini plus haut. Le processus étant markovien fort, la quantité $x_s(\tau_1, \omega)$ est connue. La discontinuité entraîne l'existence d'un intervalle $]\tau_1, \tau_1+\delta[$ tel que $x_s(t, \omega) = x_s(\tau_1, \omega)$ pour $t \in\,]\tau_1, \tau_1+\delta[$. Soit $\tau_2 = \sup\,[t: x_s(u, \omega) = x_s(\tau_1, \omega),\ \tau_1 \leqslant u < t]$. On vérifie immédiatement que τ_2 est aussi s-markovien. De toute évidence $\tau_1 < \tau_2$. Si l'instant τ_{k-1} est défini, l'instant τ_k est donné par

$$\tau_k = \sup\,[t: x_s(u, \omega) = x_s(\tau_{k-1}, \omega),\ \tau_{k-1} \leqslant u < t].$$

Il peut s'avérer que $\tau_k = +\infty$ pour un certain k. Les τ_j ne seront pas alors définis pour $j > k$ de même que $x_s(\tau_k, \omega)$. Pour construire le processus (ses réalisations) sur l'intervalle $[s, \sup\limits_k \tau_k]$ il suffit de connaître la suite de couples $\{(\tau_k, x_s(\tau_k, \omega)),\ k = 1, 2, \ldots\}$.

THÉORÈME 2. *La suite* $\{\tau_k, x_s(\tau_k, \omega)\}$, $k = 1, 2, \ldots$, *forme généralement une chaîne de Markov homogène tronquée dans l'espace des phases* $([0, \infty[\, \times X, \mathfrak{A} \times \mathfrak{B})$, *où* $\mathfrak{A}$ *est la tribu des boréliens de* $[0, \infty[$, *dont la probabilité de passage en une transition est définie par*

$$\mathsf{P}(\tau_m < t, x_s(\tau_m, \omega) \in B \mid \tau_{m-1}, x_s(\tau_{m-1}, \omega)) = \\ = \mathsf{P}_{u,x}(\tau_1 < t, x_u(\tau_1, \omega) \in B)\,\big|_{\substack{u=\tau_{m-1},\\ x=x_s(\tau_{m-1}, \omega)}}, \quad t > s, \quad B \in \mathfrak{B}. \quad (5)$$

D é m o n s t r a t i o n. Pour démontrer le théorème il suffit de prouver que le second membre de (5) est confondu avec

$$\mathsf{P}_{s,x}(\{\tau_m \in A,\ x_s(\tau_m, \omega) \in B\} \mid \mathfrak{S}^s_{\tau_{m-1}}).$$

Or, le processus étant discontinu, on a

$$\{\tau_m < t,\ x_s(\tau_m, \omega) \in B\} = \bigcup_n \bigcup_{\tau_{m-1}+\frac{k}{2^n}<t} \left[\bigcap_{i=1}^{k-1} \left\{x_s\left(\tau_{m-1}+\frac{i}{2^n}, \omega\right) = \right.\right.$$
$$\left.\left. = x_s(\tau_{m-1}, \omega)\right\}\right] \cap \left\{x_s\left(\tau_{m-1}+\frac{k}{2^n}, \omega\right) \in B \setminus \{x_s(\tau_{m-1}, \omega)\}\right\}.$$

En vertu de (18), § 1, pour toute variable ξ bornée $\mathfrak{S}^s_{\tau_{m-1}}$-mesurable on a

$$\mathsf{E}_{s,x}\xi \sum_{\tau_{m-1}+\frac{k}{2^n}<t} \prod_{i=1}^{k-1} \chi\left\{x_s\left(\tau_{m-1}+\frac{i}{2^n}, \omega\right) = \right.$$
$$\left. = x_s(\tau_{m-1}, \omega)\right\} \chi\left\{x_s\left(\tau_{m-1}+\frac{k}{2^n}, \omega\right) \in B \setminus \{x_s(\tau_{m-1}, \omega)\}\right\} =$$
$$= \mathsf{E}_{s,x}\xi \mathsf{E}_{\tau_{m-1}, x_s(\tau_{m-1}(\omega))} \sum_{\tau_{m-1}+\frac{k}{2^n}<t} \prod_{i=1}^{k-1} \chi\left\{x_{\tau_{m-1}}\left(\tau_{m-1}+\frac{i}{2^n}, \omega\right) = \right.$$
$$\left. = x_s\{\tau_{m-1}, \omega\}\right\} \chi\left\{x_{\tau_{m-1}}\left(\tau_{m-1}+\frac{k}{2^n}, \omega\right) \in B \setminus \{x_s(\tau_{m-1}, \omega)\}\right\}. \quad (6)$$

En passant à la limite pour $n \to \infty$ à gauche, on trouve

$$\mathsf{E}_{s,x}\xi\chi\{\tau_m < t,\ x_s(\tau_m, \omega) \in B\}.$$

Si l'on remarque que

$$\lim_{n\to\infty} \mathsf{E}_{u,x} \sum_{u+\frac{k}{2^n}<t} \prod_{i=1}^{k-1} \chi\left(x_u\left(u+\frac{i}{2^n}, \omega\right) = x\right\} \times$$
$$\times \chi\left\{x_u\left(u+\frac{k}{2^n}, \omega\right) \in B \setminus \{x\}\right\} =$$
$$= \mathsf{E}_{u,x}\chi\{\tau_1 < t,\ x_u(\tau_1, \omega) \in B\} = \mathsf{P}_{u,x}\{\tau_1 < t,\ x_u(\tau_1, \omega) \in B\},$$

on obtient en définitive

$$\mathsf{E}_{s,x}\xi\chi\{\tau_m < t,\ x_s(\tau_m, \omega) \in B\} =$$
$$= \mathsf{E}_{s,x}\xi\,[\mathsf{P}_{u,x}\{\tau_1 < t,\ x_u(\tau_1, \omega) \in B\}]_{\substack{u=\tau_{m-1} \\ x=x_s(\tau_{m-1}, \omega)}}. \quad \blacksquare$$

Un processus discontinu est dit *régulier* si $\sup \tau_k = +\infty$ avec la probabilité $\mathsf{P}_{s,\,x} = 1$ quels que soient s et x. Voici une condition suffisante de régularité.

LEMME 2. *Si existe* $\delta > 0$ *tel que*

$$\mathsf{P}_{s,\,x}\{\tau_1 > s+\delta\} > \delta, \qquad \forall s,\ \forall x,$$

alors le processus est régulier.

D é m o n s t r a t i o n. Le théorème 2 nous dit que pour tous les n

$$\mathsf{P}\{\tau_n - \tau_{n-1} > \delta \mid \mathfrak{S}^s_{\tau_{n-1}}\} = \mathsf{P}_{u,\,x}\{\tau_1 > u+\delta\}\big|_{\substack{u=\tau_{n-1}\\ x=x_s(\tau_{n-1},\,\omega)}} > \delta. \qquad (7)$$

Posons $\zeta_n = \tau_n - \tau_{n-1}$. De (7) il résulte que

$$\mathsf{P}\{\zeta_n > \delta \mid \zeta_1, \ldots, \zeta_{n-1}\} \geqslant \delta, \qquad \mathsf{P}\{\zeta_n \leqslant \delta \mid \zeta_1, \ldots, \zeta_{n-1}\} \leqslant (1-\delta).$$

Pour chacun des événements $A_{i_1,\ldots,i_m} = \bigcap\limits_{k=1}^{m} \{\zeta_{i_k} \leqslant \delta\}$ on a la majoration

$$\mathsf{P}(A_{i_1,\ldots,i_m}) \leqslant (1-\delta)^m.$$

Ensuite, l'événement $\{\zeta_1 + \ldots + \zeta_n \leqslant k\delta\}$ implique l'un des événements $\{A_{i_1,\ldots,i_{n-k}},\ 1 \leqslant i_1 < \ldots < i_{n-k} \leqslant n\}$. Donc

$$\mathsf{P}\{\zeta_1 + \ldots + \zeta_n \leqslant k\delta\} \leqslant C_n^k (1-\delta)^{n-k} \to 0$$

pour $n \to \infty$ quel que soit k. Par suite $\mathsf{P}_{s,\,x}\{\tau_n \leqslant t\} \to 0$ pour $n \to \infty$. ■

Déduisons une équation intégrale pour les probabilités de passage d'un processus discontinu. Notons $\pi(s, x, A, B) = \mathsf{P}_{s,\,x} \times \times (\tau_1 \in A,\ x_s(\tau_1, \omega) \in B)$ la fonction de répartition conjointe de τ_1 et $x_s(\tau_1, \omega)$. Il vient

$$\begin{aligned}
\mathsf{P}(s, x, t, B) &= \mathsf{P}_{s,\,x}(\{x_s(t, \omega) \in B\}) = \\
&= \mathsf{P}_{s,\,x}(\{x_s(t, \omega) \in B\} \cap \{\tau \leqslant t\}) + \mathsf{P}_{s,\,x}(\{x_s(t, \omega) \in B\} \cap \\
&\quad \cap \{\tau > t\}) = \chi_B(x)\, \mathsf{E}_{s,\,x}\chi_{\{\tau > t\}} + \mathsf{E}_{s,\,x}\chi_{\{\tau \leqslant t\}}\chi_B(x_s(t, \omega)) = \\
&= \chi_B(x)\, \mathsf{E}_{s,\,x}\chi_{\{\tau > t\}} + \mathsf{E}_{s,\,x}\chi_{\{\tau \leqslant t\}}\mathsf{E}_{s,\,x}(\chi_B(x_s(t, \omega)) \mid \mathfrak{S}^s_\tau) = \\
&= \chi_B(x)\, \mathsf{E}_{s,\,x}\chi_{\{\tau > t\}} + \mathsf{E}_{s,\,x}\chi_{\{\tau \leqslant t\}}\mathsf{E}_{\tau,\,x_s(\tau,\,\omega)}\chi_B(x_s(t, \omega)) = \\
&= \chi_B(x)\, \mathsf{E}_{s,\,x}\chi_{\{\tau > t\}} + \mathsf{E}_{s,\,x}\chi_{\{\tau \leqslant t\}}\mathsf{P}(\tau, x_s(\tau, \omega), t, B).
\end{aligned}$$

Par suite,

$$\mathsf{P}(s, x, t, B) = \chi_B(x)\, \pi(s, x, [t, \infty[, X) + \int_s^t \int \pi(s, x, du, dy)\, \mathsf{P}(u, y, t, B). \qquad (8)$$

Nous allons étudier maintenant un processus discontinu homogène. La probabilité de passage $\mathsf{P}(s, x, t, B)$ d'un tel processus ne dépend que de la différence $t - s$: on la désignera par $\mathsf{P}(t - s, x, B)$. La condition de continuité stochastique devient maintenant

$$\lim_{t \downarrow 0} \mathsf{P}(t, x, \{x\}) = 1. \tag{9}$$

Le théorème 1 nous apprend qu'une condition nécessaire et suffisante pour qu'un processus homogène séparable soit en escalier est que

$$\varliminf_{h \downarrow 0} \frac{1 - \mathsf{P}(h, x, \{x\})}{h} < \infty. \tag{10}$$

Si cette condition est réalisée, alors τ_1, s-instant de sortie de l'état initial du processus $x_s(t, \omega)$, admet une répartition exponentielle, c'est-à-dire existe $\lambda(x)$ tel que

$$\mathsf{P}_{s, x}\{\tau_1 > t + s\} = \exp\{-\lambda(x) t\}.$$

La condition nécessaire (10) est une conséquence de ce que

$$\varliminf_{n \to \infty} \sum \mathsf{P}(t_{k-1}^{(n)}, x, t_k^{(n)}, X \setminus \{x\}) =$$

$$= \varliminf \left(1 - \mathsf{P}\left(\frac{\delta}{n}, x, X \setminus \{x\}\right)\right) = \delta \varliminf_{h \downarrow 0} \frac{1 - \mathsf{P}(h, x\{x\})}{h}$$

pour $s = t_0^{(n)} < t_1^{(n)} < \ldots < t_n^{(n)} = s + \delta$, $h = t_k^{(n)} - t_{k-1}^{(n)} = \delta/n$. A partir de cette expression on peut établir la condition suffisante (10), puisque pour ensemble de séparabilité on peut prendre l'ensemble des nombres binaires. Ensuite

$$\mathsf{P}_{s, x}\{\tau_1 < t + s\} = 1 - \lim_{n \to \infty} \left(\mathsf{P}\left(\frac{1}{2^n}, x, \{x\}\right)\right)^{k_n} =$$

$$= 1 - \lim_{n \to \infty} \left(1 + \left[\mathsf{P}\left(\frac{1}{2^n}, x, \{x\}\right) - 1\right]\right)^{k_n} =$$

$$= 1 - \exp\left\{\varliminf_{n \to \infty} \frac{k_n}{2^n}\left(\mathsf{P}\left(\frac{1}{2^n}, x, \{x\}\right) - 1\right)\right\} = 1 - \exp\{-t\lambda(x)\},$$

où

$$\lambda(x) = \varliminf_{h \downarrow 0} \frac{1 - \mathsf{P}(h, x, \{x\})}{h}, \quad k_n = [t2^n - 1] + 1,$$

$[\cdot]$ est la partie entière d'un nombre.

Supposons maintenant que le processus $x_s(t, \omega)$ est discontinu. Il vient

$$\mathsf{P}_{s,x}\{\tau_1 < t+s,\ x_s(\tau_1, \omega) \in B\} = \lim_{n\to\infty} \sum_{k<t2^n} \mathsf{P}_{s,x}\left\{x_s\left(s+\frac{i}{2^n},\ \omega\right) = \right.$$

$$\left. = x,\ i=1, \ldots, k-1,\ x_s\left(s+\frac{k}{2^n},\ \omega\right) \in B \setminus \{x\}\right\} =$$

$$= \lim_{n\to\infty} \sum_{k<t2^n} \left[\mathsf{P}\left(\frac{1}{2^n},\ x, \{x\}\right)\right]^{k-1} \mathsf{P}\left(\frac{1}{2^n},\ x,\ B \setminus \{x\}\right) =$$

$$= \lim_{n\to\infty} \mathsf{P}\left(\frac{1}{2^n},\ x,\ B \setminus \{x\}\right) \frac{1-\mathsf{P}^{k_n}\left(\frac{1}{2^n},\ x, \{x\}\right)}{1-\mathsf{P}\left(\frac{1}{2^n},\ x, \{x\}\right)};$$

cette limite existe. Par suite existe

$$\lim_{n\to\infty} \frac{\mathsf{P}\left(\frac{1}{2^n},\ x,\ B\setminus\{x\}\right)}{1-\mathsf{P}\left(\frac{1}{2^n},\ x, \{x\}\right)} = \lim_{n\to\infty} \frac{\mathsf{P}\left(\frac{1}{2^n},\ x,\ B\setminus\{x\}\right)}{\mathsf{P}\left(\frac{1}{2^n},\ x,\ X-\{x\}\right)}.$$

Si l'on désigne cette limite par $\pi(x, B)$, on obtient

$$\mathsf{P}_{s,x}\{\tau_1 < t+s,\ x_s(\tau_1, \omega) \in B\} = \pi(x, B)\left(1-e^{-t\lambda(x)}\right). \qquad (11)$$

Donc τ_1 et $x_s(\tau_1, \omega)$ sont indépendants l'un de l'autre.

On remarquera que $\frac{1}{\lambda(x)} = \mathsf{E}_{s,x}[\tau_1 - s]$ est la durée moyenne de séjour dans l'état x. Si $\lambda(x) = 0$, $\mathsf{P}_{s,x}\{\tau_1 > t+s\} = 1$ pour tous les s, c'est-à-dire le processus ne quittera jamais cet état pourvu qu'il y soit à l'instant considéré. Un tel état est dit *absorbant.*

Etudions la quantité $\{\tau_n,\ x_s(\tau_n, \omega)\}$ dans le cas homogène. La formule (11) permet grâce au théorème 2 d'obtenir une proposition plus forte que le théorème 2 en question.

THÉORÈME 3. *Si un processus ne présente pas d'états absorbants, la suite* $\{x_n(\omega) = x_s(\tau_n, \omega),\ n = 1, 2, \ldots\}$ *forme une chaîne de Markov homogène de probabilité de passage* $\pi(x, B)$:

$$\mathsf{P}\{x_n(\omega) \in B \mid x_{n-1}(\omega)\}\big|_x = \pi(x_{n-1}(\omega),\ B),$$

les quantités $\zeta_1 = \tau_1 - s$, $\zeta_2 = \tau_2 - \tau_1, \ldots$ *sont conditionnellement indépendantes si sont donnés* $\{x_n(\omega),\ n = 1, 2, \ldots\}$ *et ont une répartition exponentielle*:

$$\mathsf{P}\{\zeta_k > t \mid x_{k-1}(\omega)\} = \exp\{-t\lambda(x_{k-1}(\omega))\}.$$

D é m o n s t r a t i o n. Le théorème 2 et la formule (11) donnent

$$\mathsf{P}_{s,x}\{\zeta_k > t,\ x_k(\omega)\in B \mid \zeta_1, \ldots, \zeta_{k-1},\ x_1(\omega), \ldots, x_{k-1}(\omega)\} =$$
$$= \mathsf{P}_{u,x}\{\tau_1 > t+u,\ x_u(\tau_1(\omega))\in B\}\big|_{\substack{u=\tau_{k-1}\\ x=x_{k-1}(\omega)}} =$$
$$= \pi(x_{k-1}(\omega), B)\exp\{-t\lambda(x_{k-1}(\omega))\}.$$

Donc la répartition conjointe de $\zeta_1, \ldots, \zeta_k$, $x_1(\omega), \ldots, x_k(\omega)$ est définie par

$$\mathsf{P}_{s,x}\{\zeta_1 > t_1,\ \zeta_2 > t_2, \ldots, \zeta_k > t_k,\ x_1(\omega)\in B_1, \ldots, x_k(\omega)\in B_k\} =$$
$$= \int_{B_1}\pi(x, dx_1)\ldots\int_{B_k}\pi(x_{k-1}, dx_k)\exp\left\{-\sum_{i=1}^{k} t_i\lambda(x_{i-1}(\omega))\right\} \qquad (12)$$

$(x_0(\omega)=x)$. En faisant $t_i=0$ on obtient

$$\mathsf{P}_{s,x}\{x_1(\omega)\in B_1, \ldots, x_k(\omega)\in B_k\} = \int_{B_1}\pi(x, dx_1)\ldots\int_{B_k}\pi(x_{k-1}, dx_k),$$

d'où il suit que $x_k(\omega)$ forment une chaîne de Markov homogène de probabilité de passage $\pi(x, B)$. De (12) on peut déduire la répartition conditionnelle de $\zeta_1, \ldots, \zeta_k$:

$$\mathsf{P}_{s,x}\{\zeta_1 > t_1, \ldots, \zeta_k > t_k \mid x_1(\omega), \ldots, x_k(\omega)\} =$$
$$= \exp\left\{-\sum_{i=1}^{k} t_i\lambda(x_{i-1}(\omega))\right\}. \ \blacksquare$$

La chaîne de Markov $\{x_k(\omega)\}$ est dite *chaîne de Markov emboîtée* pour le processus en question.

REMARQUE. Si x est un état absorbant, alors $\zeta_1 = +\infty$ et $x_1(\omega)$ n'est pas définie. Pour un tel x posons $\pi(x, B) = \chi_B(x)$ et supposons que $x_k(\omega) = x_{k-1}(\omega)$, $\zeta_k = +\infty$ si $x_{k-1}(\omega)$ tombe dans un état absorbant. Alors $\{x_k(\omega)\}$ sera encore une chaîne de Markov homogène de probabilité de passage $\pi(x, B)$.

Si l'on met l'expression (12) sous la forme

$$\mathsf{P}_{s,x}\{\zeta_1 < t_1, \ldots, \zeta_k < t_k,\ x_1(\omega)\in B_1, \ldots, x_k(\omega)\in B_k\} =$$
$$= \int_{B_1}\pi(x, dx_1)\ldots\int_{B_k}\pi(x_{k-1}, dx_k)\prod_{i=1}^{k}(1-\exp\{-t_i\lambda(x_{i-1}(\omega))\}),$$

elle sera vraie dans le cas d'états absorbants.

A noter encore que dans le cas homogène, l'équation (8) s'écrit

$$\mathsf{P}(t, x, B) = \chi_B(x) \exp\{-t\lambda(x)\} + \lambda(x) \int_0^t e^{-s\lambda(x)} ds \int \pi(x, dy) \mathsf{P}(t-s, y, B). \quad (13)$$

Si $\lambda(x)$ est bornée, on établit l'existence et l'unicité de la solution de (13) par la méthode des approximations successives. Moyennant cette condition, il résulte du lemme 2 que le processus sera régulier puisque

$$\mathsf{P}_{s,x}\{\tau_1 > s + \delta\} = e^{-\delta\lambda(x)} \geqslant e^{-\delta c}.$$

Voici une condition générale de régularité du processus.

THÉORÈME 4. *Une condition nécessaire et suffisante pour qu'un processus markovien homogène soit régulier est que*

$$\mathsf{P}_{s,x}\left\{\sum_{k=1}^{\infty} \frac{1}{\lambda(x_k(\omega))} = +\infty\right\} = 1, \quad \forall s, x \quad (14)$$

où $x_k(\omega)$ est une chaîne de Markov emboîtée.

Démonstration. Comme

$$\mathsf{E}_{s,x}(e^{-\lambda\zeta_k} \mid x_{k-1}(\omega)) = \int_0^{\infty} e^{-\lambda t}\lambda(x_{k-1}(\omega))\, e^{-t\lambda(x_{k-1}(\omega))} dt = \frac{\lambda(x_{k-1}(\omega))}{\lambda + \lambda(x_{k-1}(\omega))},$$

on a

$$\mathsf{E}_{s,x}(e^{-\lambda\tau_n} \mid x_1(\omega), \ldots, x_n(\omega)) = \prod_{k=1}^{n} \frac{\lambda(x_{k-1}(\omega))}{\lambda + \lambda(x_{k-1}(\omega))} e^{-\lambda s}. \quad (15)$$

Donc si $\tau^* = \sup_n \tau_n$, il vient

$$\mathsf{E}_{s,x} e^{-\lambda\tau^*} = e^{-\lambda s} \mathsf{E}_{s,x} \prod_{k=1}^{\infty} \frac{\lambda(x_{k-1}(\omega))}{\lambda + \lambda(x_{k-1}(\omega))}$$

($e^{-\infty} = 0$). En passant à la limite pour $\lambda \downarrow 0$, on obtient

$$\mathsf{E}_{s,x}\chi_{\{\tau^* < \infty\}} = \mathsf{E}_{s,x} \lim_{\lambda \downarrow 0} \prod_{k=1}^{\infty} \frac{\lambda(x_{k-1}(\omega))}{\lambda + \lambda(x_{k-1}(\omega))}. \quad (16)$$

Il est aisé de voir que

$$\lim_{\lambda\downarrow 0}\prod_{k=1}^{\infty}\frac{\lambda(x_{k-1}(\omega))}{\lambda+\lambda(x_{k-1}(\omega))}=\begin{cases}1 & \text{si } \sum\limits_{k=1}^{\infty}\frac{1}{\lambda(x_{k-1}(\omega))}<\infty,\\ 0 & \text{si } \sum\limits_{k=1}^{\infty}\frac{1}{\lambda(x_{k-1}(\omega))}=+\infty.\end{cases}$$

Par suite

$$\mathsf{P}_{s,x}\{\tau^*<\infty\}=\mathsf{P}_{s,x}\left\{\sum_{k=1}^{\infty}\frac{1}{\lambda(x_{k-1}(\omega))}<\infty\right\}. \quad \blacksquare \qquad (17)$$

§ 3. Processus homogènes à nombre dénombrable d'états

Dans ce paragraphe nous allons étudier des processus homogènes dans un espace des phases dénombrable X. Il semble naturel d'identifier X à l'espace des nombres naturels N. Si une tribu $\mathfrak{B}$ contient tous les ensembles à un seul point, elle contiendra toutes les parties de N, puisque toute partie de N est l'union d'un nombre au plus dénombrable d'ensembles à un point. Dans cette optique il suffit de connaître les probabilités de passage $\mathsf{P}(t, x, B)$ dans le cas seulement où B est à un seul point. On utilise généralement la notation

$$\mathsf{P}(t, i, \{j\}) = p_{ij}(t).$$

Les fonctions $p_{ij}(t)$ sont appelées aussi probabilités de passage. Des propriétés générales des probabilités de passage (cf. chapitre I, § 4) il suit:

1) $p_{ij}(t)\geqslant 0,\ t\geqslant 0,\ p_{ij}(0)=\delta_{ij}\ (\delta_{ij}=1,\ i=j;\ \delta_{ij}=0,\ i\neq j)$;

2) $\sum\limits_{j\in N} p_{ij}(t)=1$;

3) $p_{ij}(t+s)=\sum\limits_{k\in N} p_{ik}(t)\,p_{kj}(s)$ pour $t>0,\ s>0$

(équation de Chapman-Kolmogorov). On supposera encore que le processus est stochastiquement continu, c'est-à-dire

4) $\lim\limits_{t\downarrow 0} p_{ij}(t)=\delta_{ij}$.

Parfois on remplace la condition 2) par une moins forte:

$$\sum_{j\in N} p_{ij}(t)\leqslant 1.$$

Le cas $\sum\limits_{j\in N} p_{ij}(t)<1$ admet l'interprétation suivante: le système qui se trouve à un certain instant dans l'état i avec une probabilité

positive égale à $1 - \sum_{j \in N} p_{ij}(t)$, au bout d'un temps t s'absente de l'espace des phases. En d'autres termes, les points de l'espace des phases ne suffisent pas à décrire tous les états du système. On conviendra d'appeler *impropres* les processus markoviens de cette nature. On remarquera aussitôt qu'en ajoutant un certain ensemble de points à l'espace des phases on peut sans changer pour autant les probabilités de passage transformer un processus markovien impropre en processus markovien au sens propre. Le plus simple est d'adjoindre l'état « absorbant infiniment éloigné ∞ » à l'espace des phases. Posons

$$N^* = N \cup \{\infty\}, \quad p_{i\infty} = 1 - \sum_{j \in N} p_{ij}(t),$$

$$p_{\infty i}(t) = 0, \quad i \in N, \quad p_{\infty\infty}(t) = 1.$$

Il est immédiat de s'assurer que l'ensemble des probabilités de passage

$$\{p_{ij}(u)\}, \quad i, j \in N^*,$$

forme un processus markovien au sens propre. Pour le prouver il suffit de vérifier que (3) est réalisée. On a

$$p_{ij}(t+s) = \sum_{\alpha \in N} p_{i\alpha}(t)\, p_{\alpha j}(s) = \sum_{\alpha \in N} p_{i\alpha}(t)\, p_{\alpha j}(s) + p_{i\infty}(t)\, p_{\infty j}(s) =$$
$$= \sum_{\alpha \in N^*} p_{i\alpha}(t)\, p_{\alpha j}(s), \quad i, j \in N,$$

$$p_{\infty j}(t+s) = 0 = \sum_{\alpha \in N} p_{\infty\alpha}(t)\, p_{\alpha j}(s) + p_{\infty\infty}(t)\, p_{\infty j}(s) =$$
$$= \sum_{\alpha \in N^*} p_{\infty\alpha}(t)\, p_{\alpha j}(s), \quad j \in N,$$

$$p_{\infty\infty}(t+s) = 1 = \sum_{\alpha \in N} p_{\infty\alpha}(t)\, p_{\alpha\infty}(s) + p_{\infty\infty}(t)\, p_{\infty\infty}(s) =$$
$$= \sum_{\alpha \in N^*} p_{\infty\alpha}(t)\, p_{\alpha\infty}(s),$$

$$p_{i\infty}(t+s) = 1 - \sum_{\alpha \in N} p_{i\alpha}(t+s) = 1 - \sum_{\alpha \in N} \sum_{\beta \in N} p_{i\beta}(t)\, p_{\beta\alpha}(s) =$$
$$= 1 - \sum_{\beta \in N} p_{i\beta}(t) \sum_{\alpha \in N} p_{\beta\alpha}(s) = 1 - \sum_{\beta \in N} p_{i\beta}(t)\,(1 - p_{\beta\infty}(s)) =$$
$$= p_{i\infty}(t) + \sum_{\beta \in N} p_{i\beta}(t)\, p_{\beta\infty}(s) = p_{i\infty}(t)\, p_{\infty\infty}(s) +$$
$$+ \sum_{\beta \in N} p_{i\beta}(t)\, p_{\beta\infty}(s) = \sum_{\beta \in N^*} p_{i\beta}(t)\, p_{\beta\infty}(s).$$

Donc, en fait les processus markoviens impropres n'élargissent pas la classe des processus markoviens. En particulier, les probabilités de passage des processus markoviens homogènes possèdent les mêmes propriétés générales que celle des impropres. Bien plus, pour ces derniers la quantité

$$p_{i\infty}(t) = 1 - \sum_{\alpha \in N} p_{i\alpha}(t)$$

possède les mêmes propriétés que la probabilité de passage $p_{i\alpha}(t)$, $\alpha \in N$. On remarquera que $p_{i\infty}(t)$ est une fonction monotone non décroissante. En effet, nous venons de voir que

$$p_{i\infty}(t+s) = p_{i\infty}(t) + \sum_{\beta \in N} p_{i\beta}(t)\, p_{\beta\infty}(s) \geqslant p_{i\infty}(t), \quad s > 0.$$

D'après ce qui précède nous limiterons notre étude aux processus markoviens au sens propre.

Les probabilités de passage d'un processus stochastiquement continu sont uniformément continues pour $t \geqslant 0$: pour $s > 0$

$$|p_{ij}(t+s) - p_{ij}(t)| \leqslant \sum_{k \in N} |p_{ik}(s) - \delta_{ik}|\, p_{kj}(t) \leqslant$$

$$\leqslant 1 - p_{ii}(s) + \sum_{k \neq i} p_{ik}(s) = 2\,(1 - p_{ii}(s)),$$

donc

$$|p_{ij}(t_1) - p_{ij}(t_2)| \leqslant 2\,(1 - p_{ii}(|t_1 - t_2|)).$$

On voit que $p_{ij}(t)$ sont uniformément continues en j.

Traitons la dérivabilité des fonctions $p_{ij}(t)$. Prouvons tout d'abord l'existence des dérivées à droite au point 0.

THÉORÈME 1. *Les limites finies ou infinies*

$$a_{ij} = \lim_{h \downarrow 0} \frac{p_{ij}(h) - \delta_{ij}}{h}$$

existent toujours. Si $i \neq j$, les a_{ij} sont finies; ou bien les a_{ii} sont finies, ou bien $a_{ii} = -\infty$; dans tous les cas $\sum_{j \neq i} a_{ij} \leqslant -a_{ii}$.

D é m o n s t r a t i o n. Soit $i = j$. Posons

$$s = \sup_{h > 0} \frac{1 - p_{ii}(h)}{h}$$

(éventuellement $s = +\infty$). Si $c < s$ et $\frac{1 - p_{ii}(t_0)}{t_0} > c$, alors pour $\frac{t_0}{n+1} \leqslant \tau < \frac{t_0}{n}$ et, compte tenu de l'inégalité

$$p_{ii}(t+h) \geqslant p_{ii}(t)\, p_{ii}(h),$$

on aura

$$c < \frac{1}{t_0}[1-[p_{ii}(\tau)]^n p_{ii}(t_0-n\tau)] <$$

$$< \frac{1-[p_{ii}(\tau)]^n+(1-p_{ii}(t_0-n\tau))}{n\tau} < \frac{[1-p_{ii}(\tau)]^n}{n\tau} + \frac{1-p_{ii}(t_0-n\tau)}{t_0}.$$

Comme $p_{ii}(t_0 - n\tau) \to 1$ pour $\tau \to 0$ et $t_0 - n\tau \to 0$, il vient

$$\varliminf_{\tau\downarrow 0} \frac{1-p_{ii}(\tau)}{\tau} \geqslant c \quad \forall c < s.$$

Vu que

$$\varlimsup_{\tau\downarrow 0} \frac{1-p_{ii}(\tau)}{\tau} \leqslant s,$$

les deux dernières relations entraînent

$$\lim_{\tau\downarrow 0} \frac{1-p_{ii}(\tau)}{\tau} = \sup_h \frac{1-p_{ii}(h)}{h}.$$

Supposons $i \neq j$. Choisissons δ tel que $p_{ii}(s) > c$, $p_{jj}(s) > c$, où $1/2 < c < 1$, pour $0 < s \leqslant nh < \delta$. Soit $\{\xi_k, k = 0, 1, \ldots\}$ chaîne de Markov d'espace des phases N et de probabilité de passage en une transition $p_{ij} = p_{ij}(h)$. On a

$$p_{ij}(nh) = \mathsf{P}\{\xi_n = j \mid \xi_0 = i\} \geqslant$$

$$\geqslant \sum_{r=0}^{n-1} \mathsf{P}\{\xi_1 \neq j, \ldots, \xi_{r-1} \neq j, \xi_r = i \mid \xi_0 = i\} p_{ij} \mathsf{P}\{\xi_n = j \mid \xi_{r+1} = j\} \geqslant$$

$$\geqslant c p_{ij} \sum_{r=0}^{n-1} \mathsf{P}\{\xi_1 \neq j, \ldots, \xi_{r-1} \neq j, \xi_r = i \mid \xi_0 = i\}.$$

Or

$$\mathsf{P}\{\xi_1 \neq j, \ldots, \xi_{r-1} \neq j, \xi_r = i \mid \xi_0 = i\} = \mathsf{P}\{\xi_r = i \mid \xi_0 = i\} -$$

$$- \sum_{l<r} \mathsf{P}\{\xi_1 \neq j, \ldots, \xi_{l-1} \neq j, \xi_l = j \mid \xi_0 = i\} \mathsf{P}\{\xi_r = i \mid \xi_l = j\} \geqslant$$

$$\geqslant c - (1-c) \sum_{l<r} \mathsf{P}\{\xi_1 \neq j, \ldots, \xi_{l-1} \neq j, \xi_l = j \mid \xi_0 = i\} \geqslant 2c - 1.$$

Donc

$$p_{ij}(nh) \geqslant c(2c-1) n p_{ij}(h).$$

Soit $t < \delta$, $h < \delta$, $n = \left[\frac{t}{h}\right]$ (partie entière). Il vient

$$\frac{p_{ij}(h)}{h} \leqslant \frac{1}{c(2c-1)} \frac{p_{ij}\left(\left[\frac{t}{h}\right]h\right)}{\left[\frac{t}{h}\right]h}.$$

En passant à la limite pour $h \downarrow 0$, on trouve

$$\overline{\lim_{h \downarrow 0}} \frac{p_{ij}(h)}{h} \leqslant \frac{1}{c(2c-1)} \frac{p_{ij}(t)}{t} < \infty .$$

Donc

$$\overline{\lim_{h \downarrow 0}} \frac{p_{ij}(h)}{h} \leqslant \frac{1}{c(2c-1)} \underline{\lim_{t \downarrow 0}} \frac{p_{ij}(t)}{t} .$$

Or, en choisissant δ arbitrairement petit, on peut rendre c aussi proche que l'on veut de 1. Donc

$$\overline{\lim_{h \downarrow 0}} \frac{p_{ij}(h)}{h} \leqslant \underline{\lim_{t \downarrow 0}} \frac{p_{ij}(t)}{t} ,$$

$$\lim_{h \downarrow 0} \frac{p_{ij}(h)}{h} = \overline{\lim_{h \downarrow 0}} \frac{p_{ij}(h)}{h} < \infty .$$

Si $N_1 \subset N$ est un ensemble fini d'états ne contenant pas i, alors

$$\frac{1-p_{ii}(h)}{h} \geqslant \sum_{j \in N_1} \frac{1}{h} p_{ij}(h),$$

d'où

$$-a_{ii} \geqslant \sum_{j \in N_1} a_{ij}$$

et

$$-a_{ii} \geqslant \sum_{j \neq i} a_{ij} . \quad \blacksquare$$

Les quantités a_{ij} permettent de classer les états du processus. L'état i est *instantané* si $a_{ii} = -\infty$; dans le cas contraire il est *non instantané* ou *retardateur*. Un état retardateur i est régulier si

$$\sum_{j \neq i} a_{ij} = -a_{ii} .$$

Dans le cas contraire il est *non régulier*.

Soit i retardateur. On supposera que le processus est séparable. L'instant ζ de première sortie de l'état i a alors la répartition exponentielle:

$$\mathsf{P}\{\zeta > t \mid x_0(0, \omega) = i\} = \lim_{n \to \infty} \mathsf{P}\{x_0(t_{nk}, \omega) = i, \ k = 1, \ldots, n\},$$

où $0 = t_{n0} < t_{n1} < \ldots < t_{nn} = t$ et $\max(t_{n,k+1} - t_{nk}) \to 0$, les ensembles $\Lambda_n = \{t_{n1}, \ldots, t_{n,k-1}\}$ sont monotones croissants et

$\bigcup \Lambda_n = \Lambda \cap [0, t]$, où Λ est ensemble de séparabilité. Donc

$$\mathsf{P}\{\zeta > t \mid x_0(0, \omega) = i\} = \lim_{n\to\infty} \prod_{k=1}^{n} p_{ij}(t_{nk} - t_{n,k-1}) =$$

$$= \exp\left\{\lim_{n\to\infty} \sum_{k=1}^{n} \ln p_{ii}(t_{nk} - t_{n,k-1})\right\} = \exp\{a_{ii}t\},$$

puisque $\ln p_{ii}(t_{nk}-t_{n,k-1}) \sim p_{ii}(t_{nk}-t_{n,k-1})-1 \sim a_{ii}(t_{nk}-t_{n,k-1})$.

Supposons maintenant que le processus est continu à droite à l'instant ζ de première sortie de l'état régulier i. Trouvons la répartition de $x_0(\zeta, \omega)$. On a

$$\mathsf{P}\{x_0(\zeta, \omega) = j \mid x_0(0, \omega) = i\} =$$

$$= \lim_{n\to\infty} \mathsf{P}_{0,i}\left\{\bigcup_k \left\{x_0\left(\frac{1}{2^n}, \omega\right) = i, \ldots, x_0\left(\frac{k-1}{2^n}, \omega\right) = \right.\right.$$

$$\left.\left. = i,\ x_0\left(\frac{k}{2^n}, \omega\right) = j\right\}\right\} =$$

$$= \lim_{n\to\infty} \sum_{k=1}^{\infty} \left[p_{ii}\left(\frac{1}{2^n}\right)\right]^{k-1} p_{ij}\left(\frac{1}{2^n}\right) = \lim_{n\to\infty} \frac{p_{ij}\left(\frac{1}{2^n}\right)}{1 - p_{ii}\left(\frac{1}{2^n}\right)} = \frac{a_{ij}}{-a_{ii}}.$$

Telle est la signification probabiliste des coefficients a_{ij}.

Etudions maintenant la dérivabilité des probabilités de passage pour $t > 0$. Si i est un état régulier, alors pour $h > 0$

$$p_{ij}(t+h) - p_{ij}(t) = \sum_k [p_{ik}(h) - \delta_{ik}]\, p_{kj}(t) =$$

$$= (p_{ii}(h) - 1)\, p_{ij}(t) + \sum_{k\neq i} p_{ik}(h)\, p_{kj}(t).$$

Choisissons N_1, ensemble fini ne contenant pas i, tel que

$$-a_{ii} - \sum_{j\in N_1} a_{ij} < \varepsilon.$$

Si h est si petit que

$$\left|\frac{1-p_{ii}(h)}{h} + a_{ii}\right| < \varepsilon, \quad \sum_{j\in N_1} \left|\frac{p_{ij}(h)}{h} - a_{ij}\right| < \varepsilon,$$

alors

$$\sum_{j\neq i,\ j\bar{\in} N_1} \frac{p_{ij}(h)}{h} + \frac{1-p_{ii}(h)}{h} + \sum_{j\in N_1} \frac{p_{ij}(h)}{h} \leqslant$$

$$\leqslant -a_{ii} + \sum_{j\in N} a_{ij} + \left|\frac{1-p_{ii}(h)}{h} + a_{ii}\right| + \sum_{j\in N_1} \left|\frac{p_{ij}(h)}{h} - a_{ij}\right| < 3\varepsilon.$$

Donc

$$\left|\frac{p_{ij}(t+h)-p_{ij}(t)}{h}-\frac{p_{ii}(h)-1}{h}p_{ij}(t)-\sum_{k\in N_1}\frac{p_{ik}(h)}{h}p_{kj}(t)\right|\leqslant$$

$$\leqslant\sum_{\substack{k\neq i\\ k\bar{\in}N_1}}\frac{p_{ik}(h)}{h}p_{kj}(t)\leqslant 3\varepsilon.$$

Par suite

$$\left|\frac{p_{ij}(t+h)-p_{ij}(t)}{h}-a_{ii}p_{ij}(t)-\sum_{k\in N_1}a_{ik}p_{kj}(t)\right|\leqslant 5\varepsilon.$$

Comme

$$\sum_{k\neq i,\ k\bar{\in}N_1}a_{ik}p_{kj}(t)\leqslant\sum_{k\neq i,\ k\bar{\in}N_1}a_{ik}=a_{ii}-\sum_{j\in N_1}a_{ij}<\varepsilon,$$

il vient

$$\left|\frac{p_{ij}(t+h)-p_{ij}(t)}{h}-\sum a_{ik}p_{kj}(t)\right|<6\varepsilon.$$

D'où suit l'existence de la dérivée à droite $\frac{d^+p_{ij}(t)}{dt}$ et de l'égalité

$$\frac{d^+p_{ij}(t)}{dt}=\sum a_{ik}p_{kj}(t).$$

Le second membre étant continu, la dérivée le sera aussi et par suite se confond avec une dérivée ordinaire. On a donc prouvé que

$$\frac{dp_{ij}(t)}{dt}=\sum a_{ik}p_{kj}(t). \tag{1}$$

Théorème 2. *Si tous les états d'un processus sont réguliers, les probabilités de passage* $p_{ij}(t)$ $(i, j\in N)$ *sont solutions du système d'équations* (1), *appelé premier système* (*ou système inverse*) *de Kolmogorov.*

Montrons que le système d'équations (1) admet toujours une solution $p_{kj}(t)$ vérifiant les conditions 1) à 4), pourvu que les coefficients a_{ik} soient tels que:

1) $\quad -a_{ii}\geqslant 0,\quad a_{ik}\geqslant 0,\quad k\neq i;\quad$ 2) $\quad \sum_k a_{ik}=0.$

Cette solution peut être construite par la méthode des approximations successives. Pour cela mettons (1) sous la forme

$$p_{ij}(t)=\delta_{ij}e^{a_{ii}t}+\int_0^t e^{a_{ii}(t-s)}\sum_{k\neq i}a_{ik}p_{ki}(s)\,ds. \tag{2}$$

Comme

$$\frac{dp_{ii}(t)}{dt} - a_{ii}p_{ij}(t) = \sum_{k \neq i} a_{ik}p_{kj}(t),$$

il vient

$$e^{a_{ii}t}\frac{d}{dt}[p_{ij}(t)e^{-a_{ii}t}] = \sum_{k \neq i} a_{ik}p_{kj}(t),$$

$$\frac{d}{ds}[p_{ij}e^{-a_{ii}s}] = e^{-a_{ii}s}\sum_{k \neq i} a_{ik}p_{kj}(s).$$

Une intégration entre 0 et t nous donne (2).

Posons ensuite

$$p_{ij}^{(0)}(t) = \delta_{ij}e^{a_{ii}t}, \tag{3}$$

$$p_{ij}^{(n)}(t) = \delta_{ij}e^{a_{ii}t} + \int_0^t e^{a_{ii}(t-s)}\sum_{k \neq i} a_{ik}p_{kj}^{(n-1)}(s)\,ds \quad (n > 0). \tag{4}$$

Il vient

$$p_{ij}^{(1)}(t) - p_{ij}^{(0)}(t) = \int_0^t e^{a_{ii}(t-s)}\sum_{k \neq i} a_{ik}p_{kj}^{(0)}\,ds, \tag{5}$$

$$p_{ij}^{(n+1)}(t) - p_{ij}^{(n)}(t) = \int_0^t e^{a_{ii}(t-s)}\sum_{k \neq i} a_{ik}[p_{kj}^{(n)}(s) - p_{kj}^{(n-1)}(s)]\,ds. \tag{6}$$

De (5) et (6) il suit que $0 \leqslant p_{ij}^{(0)}(t) \leqslant p_{ij}^{(1)}(t) \leqslant \ldots \leqslant p_{ij}^{(n)}(t)$. Si $s_i^{(n)}(t) = \sum_j p_{ij}^{(n)}(t)$, de (3) et (4) on obtient

$$s_i^{(0)}(t) = e^{a_{ii}t} \leqslant 1,$$

$$s_i^{(n)}(t) = e^{a_{ii}t} + \int_0^t e^{a_{ii}(t-s)}\sum a_{ik}s_k^{(n-1)}(s)\,ds \leqslant$$

$$\leqslant e^{a_{ii}t} + \int_0^t e^{a_{ii}(t-s)}\sum_{k \neq i} a_{ik}\,ds = e^{a_{ii}t} - a_{ii}\int_0^t e^{a_{ii}(t-s)}\,ds = 1,$$

pourvu que $s_k^{(n-1)}(t) \leqslant 1$. Par suite $\forall n$

$$\sum_j p_{ij}^{(n)}(t) \leqslant 1.$$

Donc existent les limites

$$\bar{p}_{ij}(t) = \lim_{n\to\infty} p_{ij}^{(n)}(t). \tag{7}$$

En passant à la limite dans la relation (4) on s'assure que $\bar{p}_{ij}(t)$ vérifie (2). Montrons que $\bar{p}_{ij}(t)$ sont les plus petites solutions non négatives de (2). Soit $\tilde{p}_{ij}(t)$ solution non négative de (2). On a

$$\tilde{p}_{ij}(t) \geqslant \delta_{ij} e^{a_{ii}t} = p_{ij}^{(0)}(t).$$

Si $\tilde{p}_{ij}(t) \geqslant p_{ij}^{(n-1)}(t)$ pour un $n > 0$, alors

$$\tilde{p}_{ij}(t) \geqslant \delta_{ij} e^{a_{ii}t} + \int_0^t e^{a_{ii}(t-s)} \sum_{k\neq i} a_{ik} p_{kj}^{(n-1)}(s)\, ds = p_{ij}^{(n)}(t).$$

Donc

$$\tilde{p}_{ij}(t) \geqslant p_{ij}^{(n)}(t) \ \forall n.$$

En passant à la limite, on obtient

$$\tilde{p}_{ij}(t) \geqslant \bar{p}_{ij}(t).$$

$\bar{p}_{ij}$ est la plus petite solution. Avant d'établir l'équation de Chapman-Kolmogorov nous allons trouver une représentation pour $\bar{p}_{ij}(t)$.

En utilisant les relations de récurrence (4) on s'assure que

$$p_{ij}^{(n)}(t) = \delta_{ij} e^{a_{ii}t} + \sum_{r=1}^{n} \sum_{\substack{k_1\neq i \\ k_2\neq k_1,\ \ldots,\ k_{r-1}\neq k_{r-2}}} a_{ik_1} a_{k_1k_2} \ldots a_{k_{r-1}j} \times$$

$$\times \int\limits_{s_1+\ldots+s_r<t} \exp\{a_{ii}s_1 + a_{k_1k_1}s_2 + \ldots a_{k_{r-1}k_{r-1}}s_r +$$

$$+ a_{jj}(t - s_1 - \ldots - s_r)\}\, ds_1 \ldots ds_r$$

(on admet que $k_0 = i$). Donc

$$\bar{p}_{ij}(t) = \delta_{ij} e^{a_{ii}t} + \sum_{r=1}^{\infty} \sum_{\substack{k_1\neq i \\ k_2\neq k_1,\ \ldots,\ k_{r-1}\neq k_{r-2}}} a_{ik_1} \ldots a_{k_{r-1}j} \times$$

$$\times \int\limits_{s_1+s_2+\ldots+s_r<t} \exp\{a_{ii}s_1 + \ldots + a_{k_{r-1}k_{r-1}}s_r +$$

$$+ a_{jj}(t - s_1 - \ldots - s_r)\}\, ds_1 \ldots ds_r. \tag{8}$$

Notons $\Pi(t)$ la matrice $\|p_{ij}(t)\|$; soit

$$\pi_{ij}=\begin{cases} -\dfrac{a_{ij}}{a_{ii}}, & i\neq j, \\ 0, & i=j, \end{cases} \qquad \lambda_i=-a_{ii},\ \Pi=\|\pi_{ij}\|,\ \Lambda=\|\lambda_i\delta_{ij}\|.$$

(8) s'écrit alors

$$\Pi(t)=Ie^{-t\Lambda}+$$
$$+\sum_{r=1}^{\infty}\int\limits_{s_1+\ldots+s_r<t} e^{-s_1\Lambda}\Lambda\Pi e^{-s_2\Lambda}\Lambda\Pi\ldots e^{-(t-s_1-\ldots-s_r)\Lambda}\,ds_1\ldots ds_r \tag{9}$$

(I est la matrice unitaire).

La formule (9) admet une interprétation probabiliste simple. Soit $\{x_n(\omega),\ n=0,\ 1,\ \ldots\}$ chaîne de Markov d'espace des phases N et de matrice des probabilités de passage en une transition Π. Etudions la suite de variables aléatoires $\zeta_1,\ \zeta_2,\ \ldots$ attachées à une chaîne de Markov et dont la répartition conjointe sachant $x_0(\omega)$, $x_1(\omega)$, ... coïncide avec celle de variables exponentielles indépendantes:

$$\mathsf{P}\{\zeta_k>t\,|\,x_0(\omega),\ x_1(\omega),\ \ldots\}=\exp\{-\lambda_{x_{k-1}(\omega)}t\}.$$

En d'autres termes, $\{x_n(\omega)\}$ est chaîne de Markov emboîtée pour le processus dont on veut construire les probabilités de passage, et $\{\zeta_1,\ \zeta_2,\ \ldots\}$ sont les durées de séjour dans les divers états. Soit

$$\bar{x}(t,\ \omega)=x_n(\omega) \text{ si } \sum_{k=1}^{n}\zeta_k\leqslant t<\sum_{k=1}^{n+1}\zeta_k \quad \left(\sum_{1}^{0}=0\right).$$

Le processus $\bar{x}(t,\ \omega)$ est défini sur $\left[0,\ \sum_{k=1}^{\infty}\zeta_k\right[$, il s'arrête à l'instant $\sum_{k=1}^{\infty}\zeta_k$ (on peut admettre qu'il est tombé dans l'état absorbant $+\infty$). Notons

$$q_{ij}^{(r)}(t)=\mathsf{P}\left\{\bar{x}(t,\ \omega)=j,\ \sum_{k=1}^{r}\zeta_k\leqslant t<\sum_{k=1}^{r+1}\zeta_k\,|\,\bar{x}(0,\ \omega)=i\right\} \tag{10}$$

(la probabilité de passer de i en j exactement en r transitions). Il vient

$$q_{ij}^{(0)}(t)=\delta_{ij}e^{-t\lambda_i},$$
$$q_{ij}^{(r)}(t)=\sum_{k_1,\ \ldots,\ k_r}\ \int\limits_{s_1+\ldots+s_r<t} e^{-\lambda_i s_1}\lambda_i\pi_{ik_1}e^{-\lambda_{k_1}s_2}\lambda_{k_1}\pi_{k_1k_2}\times\ldots$$
$$\ldots\times e^{-\lambda_{k_r}s_r}\lambda_{k_r}\pi_{k_rj}e^{-\lambda_j(t-s_1-\ldots-s_r)}\,ds_1\ldots ds_r.$$

Donc, en appelant $Q^{(r)}(t)=\|q_{ij}^{(r)}(t)\|$, (9) devient

$$\Pi(t)=\sum_{r=0}^{\infty} Q^{(r)}(t). \tag{11}$$

L'égalité (10) nous dit que

$$q_{ij}^{(r)}(t)=\sum_k \sum_{l=0}^{r} \mathsf{P}\left\{\bar{x}(t+u,\ \omega)=j,\ \bar{x}(t,\ \omega)=k,\right.$$

$$\left.\sum_{m=1}^{l} \zeta_m \leqslant t < \sum_{m=1}^{l+1} \zeta_k,\ \sum_{n=l+1}^{r} \zeta_n \leqslant u < \sum_{n=l+1}^{r+1} \mid \bar{x}(0,\ \omega)=i\right\}=$$

$$=\sum_k \sum_{l=0}^{r} q_{ik}^{(l)}(t)\, q_{kj}^{(r-l)}(u),$$

i.e.

$$Q^{(r)}(t+u)=\sum_{l=0}^{r} Q^{(l)}(t)\, Q^{(r-l)}(u).$$

Par suite

$$\Pi(t+u)=\sum_{r=0}^{\infty} Q^{(r)}(t+u)=\sum_{r=0}^{\infty} \sum_{l=0}^{r} Q^{(l)}(t)\, Q^{(r-l)}(u)=$$

$$=\sum_{\substack{r\geqslant 0\\ q\geqslant 0}} Q^{(r)}(t)\, Q^{(q)}(u)=\Pi(t)\,\Pi(u).$$

On vient donc d'établir l'équation de Chapman-Kolmogorov pour $\bar{p}_{ij}(t)$.

L'unicité de la solution du premier système de Kolmogorov vérifiant la condition initiale $p_{ij}(0)=\delta_{ij}$ est directement liée à la régularité du processus. On remarquera en effet qu'une condition suffisante d'unicité de la solution est que

$$\sum_j \bar{p}_{ij}(t)=1 \quad \forall i,\ \forall t>0 \tag{12}$$

($\bar{p}_{ij}$ est la plus petite solution). En effet, si $p_{ij}(t)$ était une autre solution non négative telle que $\sum p_{ij}(t)=1$, on aurait

$$p_{ij}(t)-\bar{p}_{ij}(t)\geqslant 0$$

et

$$\sum_j (p_{ij}(t)-\bar{p}_{ij}(t))=1-1=0.$$

Donc

$$p_{ij}(t)=\bar{p}_{ij}(t).$$

Or, de (10) il suit:

$$\sum_j \bar{p}_{ij}(t)=\sum_{r=0}^{\infty} \mathsf{P}\left\{\sum_{k=1}^{r}\zeta_k\leqslant t<\sum_{k=1}^{r+1}\zeta_k \mid \bar{x}(0,\,\omega)=i\right\}=$$

$$=\mathsf{P}\left\{\sup_r \sum_{k=1}^{r}\zeta_k>t \mid \bar{x}(0,\,\omega)=i\right\}=\mathsf{P}\left\{\sum_{k=1}^{\infty}\zeta_k>t \mid \bar{x}(0,\,\omega)=i\right\}.$$

Par suite, la condition (12) équivaut à la régularité du processus. Dans le cas d'un processus non régulier il est possible de construire d'autres solutions du premier système de Kolmogorov. Indiquons une autre méthode de construction. On se donne des probabilités quelconques p_k, $\sum p_k=1$, et on suppose que le processus prend l'état k avec la probabilité p_k à l'instant $\sum_{k=1}^{\infty}\zeta_k$ (s'il est fini). Les probabilités de passage d'un tel processus seront solutions de l'équation suivante:

$$p_{ij}(t)=\mathsf{P}\left\{x(t,\,\omega)=j,\ \sum_{k=1}^{\infty}\zeta_k>t \mid x(0,\,\omega)=i\right\}+$$

$$+\sum_l \int_0^t \mathsf{P}\left\{\sum_{k=1}^{\infty}\zeta_k\in ds \mid \bar{x}(0,\,\omega)=i\right\} p_l p_{lj}(t-s) \qquad (13)$$

(on peut passer de i en j en un nombre fini de sauts, ou bien après une infinité de sauts). Le premier terme du second membre est confondu avec $\bar{p}_{ij}(t)$. Les fonctions

$$\Phi_i(s)=\mathsf{P}\left\{\sum_{k=1}^{\infty}\zeta_k<s \mid \bar{x}(0,\,\omega)=i\right\}$$

peuvent être supposées données. L'équation (13) devient

$$p_{ij}(t)=\bar{p}_{ij}(t)+\sum_l \int_0^t p_l p_{lj}(t-s)\,d\Phi_i(s). \qquad (14)$$

La solution vérifiant les conditions 1) à 4) se construit exactement comme la plus petite solution de l'équation (2).

Etudions maintenant le deuxième système (ou système direct) de Kolmogorov. Formellement il se déduit de l'équation de Chapman-Kolmogorov de la manière suivante:

$$\frac{p_{ij}(t+h)-p_{ij}(t)}{h}=\sum_k p_{ik}(t)\left[\frac{p_{kj}(h)-\delta_{kj}}{h}\right],$$

et en passant à la limite

$$\frac{dp_{ij}(t)}{dt} = \sum_k p_{ik}(t)\, a_{kj}. \tag{15}$$

L'équation (15) est commode, car on peut en déduire l'équation de la répartition conditionnelle du processus: si $\mathsf{P}\{x(0, \omega) = i\} = p_i$, alors

$$\mathsf{P}\{x(t, \omega) = j\} = \sum_i p_i p_{ij}(t).$$

En multipliant (15) par p_i et en sommant sur i on trouve

$$\frac{d}{dt}\mathsf{P}\{x(t, \omega) = j\} = \sum_k \mathsf{P}\{x(t, \omega) = k\}\, a_{kj}. \tag{16}$$

Voici encore une condition suffisante de solubilité du deuxième système de Kolmogorov.

THÉORÈME 3. *Si tous les états du processus sont réguliers et*

$$\sum_k p_{ik}(t)\, a_{kk} > -\infty \tag{17}$$

pour i donné, alors (15) *est réalisée pour tous les j.*

D é m o n s t r a t i o n. Pour démontrer le théorème 1 on a établi que

$$\frac{1 - p_{ii}(h)}{h} \leqslant -a_{ii}.$$

Donc pour $k \neq i$

$$\frac{p_{kj}(h)}{h} \leqslant \frac{1 - p_{kk}(h)}{h} \leqslant -a_{kk}.$$

On a

$$\frac{p_{ij}(t+h) - p_{ij}(t)}{h} = \sum_k p_{ik}(t)\left[\frac{p_{kj}(h) - \delta_{kj}}{h}\right].$$

La série du second membre étant majorable par rapport à h:

$$\sum_k \left| p_{ik}(t)\, \frac{p_{kj}(h) - \delta_{kj}}{h} \right| \leqslant \sum_k | p_{ik}(t)\, a_{kk} |,$$

on peut passer à la limite pour $h \downarrow 0$. L'existence de la dérivée à gauche a été établie au théorème 2. ■

On remarquera que la série de (15) est toujours convergente quelles que soient les probabilités de passage, pourvu que tous les états du processus soient réguliers. En effet,

$$\frac{p_{ij}(t+h) - p_{ij}(t)}{h} + \frac{1 - p_{jj}(h)}{h}\, p_{ij}(t) = \sum_{k \neq j} p_{ik}(t)\, \frac{p_{kj}(h)}{h};$$

en passant à la limite pour $h \downarrow 0$ et comme le théorème 2 affirme l'existence de la dérivée $\frac{d}{dt} p_{ij}(t)$, on obtient

$$\sum_{k \neq j} p_{ik}(t)\, a_{kj} \leqslant \frac{dp_{ij}(t)}{dt} - a_{jj} p_{ij}(t).$$

Montrons que la plus petite solution $\overline{p}_{ij}(t)$ du premier système de Kolmogorov est également solution du second. De la majoration obtenue il suit que sont définies les sommes

$$\sum_k \overline{p}_{ik}(t)\, a_{kj} = \sum_k \overline{p}_{ik}(t)\, \lambda_k \pi_{kj},$$

et par suite le produit de matrices

$$\Pi(u)\, \Lambda\Pi.$$

Il résulte de (9)

$$\int_0^t \Pi(u)\, \Lambda\Pi e^{-\Lambda(t-u)}\, du = \int_0^t e^{-u\Lambda} \Lambda\Pi e^{-(t-u)\Lambda}\, du +$$

$$+ \sum_{r=1}^{\infty} \int_0^t \int_{s_1 + \ldots + s_r < u} e^{-s_1 \Lambda} \Lambda\Pi \ldots e^{-(u-s_1-\ldots-s_r)\Lambda} \Lambda\Pi e^{-(t-u)\Lambda}\, du.$$

En faisant $s_{r+1} = u - s_1 - \ldots - s_r$ dans le terme de rang r de la somme, on obtient

$$\int_0^t \Pi(u)\, \Lambda\Pi e^{-(t-u)\Lambda}\, du =$$

$$= \sum_{r=1}^{\infty} \int_{s_1 + \ldots + s_r < t} e^{-s_1 \Lambda} \Pi\Lambda \ldots e^{-(t-s_1-\ldots-s_r)\Lambda}\, ds_1 \ldots ds_r =$$

$$= \Pi(t) - e^{-t\Lambda} I.$$

Cette expression s'écrit terme à terme :

$$\int_0^t \sum_{k \neq j} \overline{p}_{ik}(u)\, a_{kj} e^{-\lambda_j (t-u)}\, du = \overline{p}_{ij}(t) - e^{-t\lambda_j} \delta_{ij},$$

d'où

$$\int_0^t \sum_{k \neq j} \overline{p}_{ik}(u)\, a_{kj} e^{\lambda_j u}\, du = \overline{p}_{ij}(t)\, e^{\lambda_j t} - \delta_{ij}.$$

En dérivant par rapport à t on trouve

$$\sum_{k\neq j} \bar{p}_{ik}(t)\, a_{kj} e^{\lambda_j t} = \frac{d}{dt} p_{ij}(t)\, e^{\lambda_j t} + \lambda_j \bar{p}_{ij}(t)\, e^{\lambda_j t};$$

en substituant $\lambda_j = -a_{ij}$ et en simplifiant par $e^{\lambda_j t}$ on obtient (15).

Voici encore une condition de régularité du processus. Posons

$$g_i(\lambda) = \mathsf{E}\left(\exp\left\{-\lambda \sum_{k=1}^{\infty} \zeta_k\right\} \Big| x(0, \omega) = i\right) = \mathsf{E}_i \exp\left\{-\lambda \sum_{k=1}^{\infty} \zeta_k\right\} \tag{18}$$

(E_x désigne E_{0x}; la notation E_{0x} a été introduite au § 1). Les fonctions $g_i(\lambda)$ sont solutions du système d'équations

$$g_i(\lambda) = \mathsf{E}_i \exp\{-\lambda\zeta_1\}\, \mathsf{E}\left(\exp\left\{-\lambda \sum_{k=2}^{\infty} \zeta_k\right\} \Big| x(\zeta_1, \omega), \zeta_1\right) =$$

$$= \sum_j \mathsf{E}_i \exp\{-\lambda\zeta_1\}\, \chi_{\{x(\zeta_1, \omega)=j\}} \mathsf{E}\left(\exp\left\{-\lambda \sum_{k=2}^{\infty} \zeta_k \mid x(\zeta_1, \omega) = j\right\}\right).$$

Comme

$$\mathsf{E} \exp\left\{-\lambda \sum_{k=2}^{\infty} \zeta_k \mid x(\zeta_1, \omega) = j\right\} = g_j(\lambda),$$

il vient

$$g_i(\lambda) = \sum_j \mathsf{E}_i \exp\{-\lambda\zeta_1\}\, \chi_{\{x(\zeta_1, \omega)=j\}} g_i(\lambda) =$$

$$= \sum_j \pi_{ij} \frac{\lambda_i}{\lambda + \lambda_i} g_j(\lambda) = \sum_{j\neq i} \frac{a_{ij}}{\lambda - a_{ii}} g_j(\lambda).$$

Finalement on obtient l'équation suivante pour $g_i(\lambda)$:

$$\lambda g_i(\lambda) = \sum a_{ij} g_j(\lambda). \tag{19}$$

THÉORÈME 4. *Pour qu'un processus soit régulier il faut et il suffit que l'équation* (19) *ne possède pas de solutions bornées non nulles pour* $\lambda > 0$.

D é m o n s t r a t i o n. Si le processus est non régulier, les fonctions $g_i(\lambda)$ définies par les égalités (18) seront solutions bornées non nulles de (19). Soit maintenant le processus régulier. L'équation (19) équivaut à

$$g_i(\lambda) = \mathsf{E}_i \exp\{-\lambda\zeta_1\}\, g_{x_1(\omega)}(\lambda) \tag{20}$$

(ceci a été établi dans les raisonnements qui nous ont conduits à (19)). De (20) il suit

$$g_{x_1(\omega)}(\lambda) = \mathsf{E}_{x_1(\omega)} e^{-\lambda\zeta_2} g_{x_2(\omega)}(\lambda);$$

donc

$$g_i(\lambda) = \mathsf{E}_i \exp\{-\lambda(\zeta_1+\zeta_2)\} g_{x_2(\omega)}(\lambda).$$

En utilisant la relation

$$g_{x_k(\omega)}(\lambda) = \mathsf{E}_{x_k(\omega)} \exp\{-\lambda\zeta_{k+1}\} g_{x_{k+1}(\omega)}(\lambda),$$

on s'assure également que

$$g_i(\lambda) = \mathsf{E}_i \exp\{-\lambda(\zeta_1+\ldots+\zeta_n)\} g_{x_n(\omega)}(\lambda).$$

En passant à la limite pour $n \to \infty$ dans l'inégalité

$$|g_i(\lambda)| \leqslant \mathsf{E}_i \exp\{-\lambda(\zeta_1+\ldots+\zeta_n)\} \sup_k |g_k(\lambda)|$$

et compte tenu de la régularité du processus

$$\lim_{n\to\infty} \mathsf{E}_i \exp\{-\lambda(\zeta_1+\ldots+\zeta_n)\} = 0,$$

on trouve $g_i(\lambda) = 0$. ■

Considérons des processus sans arrêt à nombre fini d'états. On admettra que l'espace des phases est confondu avec l'ensemble $\{1, 2, \ldots, r\}$. De toute évidence, ces processus sont justiciables de tous les résultats obtenus plus haut pour les processus à nombre dénombrable d'états. (On peut ajouter à l'espace des phases une infinité d'états absorbants $\{r+1, r+2, \ldots\}$ tels que $p_{ij}(t) = 0$ pour $i \leqslant r$, $j > r$.) Si les probabilités de passage vérifient les conditions 1) à 4), tous les états du processus sont réguliers, puisque

$$-a_{ii} = \lim_{h\downarrow 0} \frac{1-p_{ii}(h)}{h} = \lim_{h\downarrow 0} \sum_{j\neq i} \frac{p_{ij}(h)}{h} = \sum_{j\neq i} a_{ij}$$

(le passage à la limite sous le signe somme est possible, puisque les termes sont en nombre fini). Par suite, les probabilités de passage sont solutions du premier système d'équations de Kolmogorov:

$$\frac{dp_{ij}(t)}{dt} = \sum_{k=1}^{r} a_{ik}p_{kj}(t), \quad i, j = 1, \ldots, r. \tag{21}$$

Comme

$$-\sum_{j=1}^{r} p_{kj}(t)\, a_j \leqslant -\sum_{j=1}^{r} a_j < \infty,$$

le théorème 3 nous dit que le second système d'équations de Kolmogorov est réalisé.

Soit $\Pi(t) = \| p_{ij}(t) \|$ la matrice des probabilités de passage. Si $A = \| a_{ij} \|$, (21) devient

$$\frac{d}{dt}\Pi(t) = A\Pi(t).$$

Comme $\Pi(0) = I$, on a

$$\Pi(t) = e^{tA}.$$

Le processus est régulier en vertu du théorème 4. En effet, la solution bornée de l'équation $\lambda g_i = \sum_j a_{ij} g_j$ satisfait l'inégalité

$$(\lambda - a_{ii}) \mid g_i \mid \leqslant \sum_{j \neq i} a_{ij} \mid g_j \mid \leqslant \max_j \mid g_j \mid \mid a_{ii} \mid .$$

Donc si i est tel que $\mid g_i \mid = \max_j \mid g_j \mid$, on a

$$(\lambda - a_{ii}) \max_j \mid g_j \mid \leqslant a_{ii} \mid \max_j \mid g_j \mid ,$$

ce qui n'est possible qu'à la condition que $\max_j \mid g_j \mid = 0$.

§ 4. Processus de naissance et de mort

C'est ainsi qu'on appelle le processus markovien homogène d'états $\{0, 1, 2, \ldots\}$ dans lequel de l'état n on ne peut passer qu'en $n - 1$ ou $n + 1$ et de 0 en 1. Tout état du processus est assimilé au nombre d'individus d'une population, le passage de n en $n + 1$, à la naissance d'un individu, le passage de n en $n - 1$, à la mort d'un individu; l'autogenèse (i.e. le passage $0 \to 1$) n'est pas exclue dans le cas général. Supposons qu'un tel processus est stochastiquement continu et à états réguliers. Les quantités

$$\lim_{h \downarrow 0} \frac{p_{ii}(h) - 1}{h} = a_{ii}$$

sont finies $\forall i = 0, 1, 2, \ldots$, et si

$$a_{ij} = \lim_{h \downarrow 0} \frac{p_{ij}(h)}{h}, \quad j \neq i,$$

la probabilité de passer de i directement en j est $-\frac{a_{ij}}{a_{ii}}$. De la définition il résulte que seuls $a_{i,\,i+1}$ et $a_{i,\,i-1}$ sont susceptibles d'être non nuls. Notons

$$a_{i,\,i+1} = \lambda_i, \quad a_{i,\,i-1} = \mu_i.$$

Il vient

$$p_{i,\,i+1}(t) = \lambda_i(t) + o(t), \quad p_{i,\,i-1}(t) = \mu_i t + o(t),$$

et par suite $\lambda_i \Delta t$ à $o(\Delta t)$ près est la probabilité de naissance d'un nouvel individu dans une population de i sujets, et $\mu_i \Delta t + o(\Delta t)$ la probabilité de mort d'un individu. Le processus étant régulier, on a $a_{ii} = -\lambda_i - \mu_i$. Le premier système d'équations de Kolmogorov s'écrit

$$\frac{d}{dt} p_{ij}(t) = -(\lambda_i + \mu_i) p_{ij}(t) + \lambda_i p_{i+1,\,j}(t) + \mu_i p_{i-1,\,j}(t), \mu_0 = 0, \quad (1)$$

et le second

$$\frac{d}{dt} p_{ij}(t) = -(\lambda_j + \mu_j) p_{ij}(t) + \lambda_{j-1} p_{i,\, j-1}(t) + \mu_{j+1} p_{i,\, j+1}(t) \qquad (2)$$

(pour $j = 0$ on pose $\lambda_{-1} = 0$). Pour que le premier système ait lieu on admettra que le processus s'arrête à la première accumulation de sauts. De (2) on obtient l'équation des probabilités inconditionnelles

$$\frac{d}{dt} p_j(t) = -(\lambda_j + \mu_j) p_j(t) + \lambda_{j-1} p_{j-1}(t) + \mu_{j+1} p_{j+1}(t) \qquad (3)$$

($p_j(0)$, la probabilité que le système se trouve dans l'état j à l'instant initial, est supposée connue). Etudions le problème important de l'existence d'une répartition stationnaire du processus, c'est-à-dire d'une répartition initiale des probabilités telle que $p_j(t)$ soient constantes. Soient $p_j(t) = p_j$. De (3) on déduit

$$-(\lambda_j + \mu_j) p_j + \lambda_{j-1} p_{j-1} + \mu_{j+1} p_{j+1} = 0. \qquad (4)$$

On admet que $\mu_j > 0$ pour $j > 0$. En faisant $j = 0$ dans (4) on trouve

$$-\lambda_0 p_0 + \mu_1 p_1 = 0, \quad p_1 = \frac{\lambda_0}{\mu_1} p_0;$$

pour $j = 1$ on a

$$\mu_2 p_2 = (\lambda_1 + \mu_1) p_1 - \lambda_0 p_0 = \frac{\lambda_1 \lambda_0}{\mu_1} p_0, \quad p_2 = \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2} p_0.$$

Un raisonnement par récurrence nous montre que

$$p_k = \frac{\lambda_0 \lambda_1 \ldots \lambda_{k-1}}{\mu_1 \mu_2 \ldots \mu_k} p_0 \qquad (5)$$

est solution du système (4). Pour que $\{p_k\}$ soit répartition de probabilités il faut et il suffit que $\sum p_k = 1$, car p_k non négatifs. Par suite, une condition nécessaire et suffisante d'une répartition stationnaire est

$$\sum_{k=1}^{\infty} \frac{\lambda_0 \lambda_1 \ldots \lambda_{k-1}}{\mu_1 \mu_2 \ldots \mu_k} < \infty; \qquad (6)$$

si cette condition est réalisée, les probabilités stationnaires sont définies par

$$p_0 = \left(1 + \sum_{k=1}^{\infty} \frac{\lambda_0 \lambda_1 \ldots \lambda_{k-1}}{\mu_1 \mu_2 \ldots \mu_k}\right)^{-1},$$

$$p_n = \frac{\lambda_0 \lambda_1 \ldots \lambda_{n-1}}{\mu_1 \mu_2 \ldots \mu_n} \left(1 + \sum_{k=1}^{\infty} \frac{\lambda_0 \ldots \lambda_{k-1}}{\mu_1 \ldots \mu_k}\right)^{-1}. \qquad (7)$$

Le processus de naissance et de mort se prête bien à la description de systèmes techniques. Voici quelques exemples.

1. *Service de machines.* Supposons que m machines sont servies par une équipe de s ouvriers.

Dès qu'une machine a un ennui mécanique, soit elle est aussitôt prise en charge par un ouvrier libre, soit elle attend son tour si les autres ouvriers sont occupés à la réparation de celles tombées en panne avant elle. Les machines sont réparées dans la « chronologie » de leurs pannes.

On postulera ce qui suit. La probabilité de panne d'une machine entre les dates t et $t + \Delta t$ ne dépend pas de t et vaut $\lambda(\Delta t) = \lambda \Delta t + o(\Delta t)$ indépendamment de l'« histoire » de son fonctionnement (i.e. de la durée de fonctionnement, du nombre de pannes et de la durée de réparation) avant l'instant t. De façon analogue, si une machine est retapée par un ouvrier, la probabilité qu'elle finisse de l'être entre les dates t et $t + \Delta t$ vaut $\mu(\Delta t) = \mu \Delta t + o(\Delta t)$ et ne dépend ni du caractère de son fonctionnement ni de la durée de service avant l'instant t. Les machines fonctionnent, tombent en panne et sont remises en état indépendamment l'une de l'autre.

On dira qu'un processus industriel est dans l'état $\mathscr{E}_k$ si à une date donnée le nombre de machines en réparation ou en attente (i.e. le nombre total de machines en panne) est égal à k. Une nouvelle panne d'une machine se traduit par un passage à l'état $\mathscr{E}_{k+1}$ et la fin d'une réparation par un passage à l'état $\mathscr{E}_{k-1}$. Nous sommes donc en présence d'un système markovien homogène à nombre fini d'états $\mathscr{E}_0, \mathscr{E}_1, \ldots, \mathscr{E}_m$. Des hypothèses faites il suit

$$\begin{aligned}
&p_{k,\,k+1}(\Delta t) = (m-k)\lambda\Delta t + o(\Delta t), && k = 0, \ldots, m-1;\\
&p_{k,\,k-1}(\Delta t) = k\mu\Delta t + o(\Delta t) && \text{pour } 1 \leqslant k \leqslant s;\\
&p_{k,\,k-1}(\Delta t) = s\mu\Delta t + o(\Delta t) && \text{pour } s \leqslant k \leqslant m;\\
&p_{k,\,k\pm r}(\Delta t) = o(\Delta t), && r \geqslant 2,
\end{aligned}$$

i.e. nous avons affaire à un processus de naissance et de mort à nombre fini de possibles. Dans les notations précédentes

$$\lambda_k = (m-k)\lambda, \quad k = 0, 1, \ldots, m;$$

$$\mu_k = \begin{cases} k\mu & \text{pour } 0 \leqslant k \leqslant s, \\ s\mu & \text{pour } s \leqslant k \leqslant m. \end{cases}$$

Il existe toujours une répartition stationnaire qui, en vertu des équations (7), est définie par

$$p_k = C_m^k \left(\frac{\lambda}{\mu}\right)^k p_0, \quad k \leqslant s,$$

$$p_k = C_m^k \frac{k(k-1)\ldots(s+1)}{s^{k-s}} \left(\frac{\lambda}{\mu}\right)^k p_0, \quad s < k \leqslant m,$$

$$p_0 = \left[1 + \sum_{k=0}^{s} C_m^k \left(\frac{\lambda}{\mu} \right)^k + \sum_{k=s+1}^{m} C_m^k \frac{k(k-1)\ldots(s+1)}{s^{k-s}} \left(\frac{\lambda}{\mu} \right)^k \right]^{-1}.$$

Ces formules peuvent servir au calcul du meilleur rapport nombre de machines — nombre d'ouvriers dans des conditions concrètes de production.

2. *Réseau téléphonique.* On a souvent affaire à des schémas analogues au précédent mais avec $m = \infty$. C'est le cas par exemple d'un interurbain de s lignes et d'un nombre pratiquement illimité d'abonnés. Les unités à servir sont ici les abonnés, les servants, les lignes. Le rang de l'état du système représente le nombre d'abonnés en attente de service. S'ils sont en nombre $> s$, ils font la queue et attendent une ligne libre. Contrairement à l'exemple précédent, on admettra que la probabilité que dans l'intervalle $(t, t + \Delta t)$ un abonné fasse une demande est égale à $\lambda(\Delta t) = \lambda \Delta t + o(\Delta t)$, et celle de plus d'un égale à $o(\Delta t)$ (ces probabilités ne dépendent pas du nombre d'abonnés ayant passé une demande avant l'instant considéré). Nous conservons le même régime de service que dans l'exemple précédent. Dans ces conditions on a

$$\lambda_k = \lambda, \ \mu_k = k\mu \text{ pour } k \leqslant s; \ \mu_k = s\mu \text{ pour } k \geqslant s.$$

Pour qu'existe une répartition stationnaire il faut et il suffit que la série

$$S = \sum_{k=0}^{s} \frac{\lambda^k}{k!\,\mu^k} + \sum_{k=s+1}^{\infty} \frac{1}{s!\,s^{k-s}} \left(\frac{\lambda}{\mu} \right)^k < \infty$$

soit convergente, i.e. $\lambda < \mu$.

La répartition stationnaire est donnée ici par:

$$p_k = \frac{1}{k!} \left(\frac{\lambda}{\mu} \right)^k \frac{1}{S}, \qquad 0 \leqslant k \leqslant s,$$

$$p_k = \frac{1}{s!\,s^{k-s}} \left(\frac{\lambda}{\mu} \right)^k \frac{1}{S}, \qquad k \geqslant s.$$

3. *Fil à tisser.* Le fil à tisser est un faisceau de fibres dont le nombre en un point donné varie suivant la longueur. Si l on suppose que la longueur d'une fibre obéit à une loi exponentielle fixe négative ne dépendant ni du nombre de fibres en un point quelconque du fil, ni de leur longueur, et que la probabilité d'apparition d'une nouvelle fibre dans l'intervalle $(t, t + \Delta t)$ est égale à $\lambda \Delta t + o(\Delta t)$ et ne dépend ni du nombre de fibres, ni de leur longueur, alors $\nu(t)$, nombre de fibres au point t, est un processus de naissance et de mort de paramètres

$$\lambda_n = \lambda, \quad \mu_n = n\mu,$$

où μ est l'inverse de la longueur moyenne d'une fibre.

Soit à déterminer la loi de répartition de $\tau_{r,\,s}$, instant de première arrivée dans l'état s à partir de l'état r $(r < s)$ d'un processus de naissance et de mort. Posons

$$F_{r,\,s}(t) = \mathsf{P}\{\tau_{r,\,s} < t\},$$

$$\varphi_{r,\,s}(z) = \int_0^\infty e^{-zt}\, dF_{r,\,s}(t) = \mathsf{E} e^{-z\tau_{r,\,s}} \qquad (z > 0).$$

En utilisant la propriété du processus d'être markovien fort, le fait que l'instant $\tau_{r,\,r+1}$ est markovien, $x(\tau_{r,\,r+1}, \omega) = r+1$ ($x(t, \omega)$ est une réalisation du processus) et enfin l'égalité (20), § 2, on trouve

$$\mathsf{P}_{0,\,r}\{\tau_{r,\,s} > t\} = \mathsf{E}_{0,\,r}\mathsf{P}\{\tau_{r,\,s} > t \mid \mathfrak{S}^0_{\tau_{r,\,r+1}}\} =$$

$$= \mathsf{E}_{0,\,r} \lim_{n\to\infty} \mathsf{P}\left\{x\left(\frac{i}{2^n} + \tau_{r,\,r+1},\ \omega\right) < s,\ \ i < t2^n \mid \mathfrak{S}^0_{\tau_{r,\,r+1}}\right\} =$$

$$= \mathsf{E}_{0,\,r} \lim_{n\to\infty} \mathsf{P}_{\tau_{r,\,r+1},\ x(\tau_{r,\,r+1},\ \omega)}\left\{x\left(\frac{i}{2^n} + \tau_{r,\,r+1},\ \omega\right) < s,\right.$$

$$\left.\frac{i}{2^n} + \tau_{r,\,r+1} < t\right\} = \mathsf{E}_{0,\,r}\left(\mathsf{P}_{0,\,r+1}\{\tau_{r+1,\,s} > t-u\}\big|_{u=\tau_{r,\,r+1}}\right) =$$

$$= \mathsf{E}_{0,\,r}\left[1 - F_{r+1,\,s}(t - \tau_{r,\,r+1})\right]. \qquad (8)$$

Donc

$$F_{r,\,s}(t) = \int_0^t F_{r+1,\,s}(t-u)\, dF_{r,\,r+1}(u).$$

Par suite,

$$\varphi_{r,\,s}(z) = \varphi_{r+1,\,s}(z)\, \varphi_{r,\,r+1}(z)$$

et

$$\varphi_{r,\,s}(z) = \varphi_{r,\,r+1}(z)\, \varphi_{r+1,\,r+2}(z) \ \cdots\ \varphi_{s-1,\,s}(z). \qquad (9)$$

Pour trouver la fonction $\varphi_{r,\,r+1}(z)$ on se servira de l'équation suivante : si ζ_r est l'instant de première sortie de l'état r $(r > 0)$, on a

$$\mathsf{P}_{0,\,r}\{\tau_{r,\,r+1} < t\} = \mathsf{P}_{0,\,r}\{\zeta_r < t,\ x(\zeta_r,\ \omega) = r+1\} +$$

$$+ \int_0^t \mathsf{P}_{0,\,r}\{\zeta_r \in ds,\ x(\zeta_r,\ \omega) = r-1\}\, \mathsf{P}_{0,\,r-1}\{\tau_{r-1,\,r+1} < t-s\}. \qquad (10)$$

Cette équation peut être déduite à l'aide de la propriété d'être markovien fort du processus exactement comme la relation (8). La trans-

formation de Laplace donne

$$\varphi_{r,\,r+1}(z)=\mathsf{E}e^{-z\zeta_r}\left[\frac{\lambda_r}{\lambda_r+\mu_r}+\frac{\mu_r}{\lambda_r+\mu_r}\,\varphi_{r-1,\,r+1}(z)\right]=$$

$$=\frac{\lambda_r+\mu_r}{z+\lambda_r+\mu_r}\left[\frac{\lambda_r}{\lambda_r+\mu_r}+\frac{\mu_r}{\lambda_r+\mu_r}\,\varphi_{r-1,\,r}(z)\,\varphi_{r,\,r+1}(z)\right]=$$

$$=\frac{\lambda_r}{z+\lambda_r+\mu_r}+\frac{\mu_r}{z+\lambda_r+\mu_r}\,\varphi_{r-1,\,r}(z)\,\varphi_{r,\,r+1}(z).$$

Donc

$$\varphi_{r,\,r+1}(z)=\frac{\lambda_r}{z+\lambda_r+\mu_r\,(1-\varphi_{r-1,\,r}(z))}. \tag{11}$$

La formule (11) permet de déterminer successivement $\varphi_{r,\,r+1}(z)$, pourvu que soit connue $\varphi_{0,\,1}(z)$. Or, par définition $\tau_{0,\,1}=\zeta_0$, donc

$$\varphi_{0,\,1}(z)=\frac{\lambda_0}{\lambda_0+z}. \tag{12}$$

Les formules (12), (11) et (9) permettent de calculer $\varphi_{r,\,s}(s)$ pour $r<s$.

Etudions maintenant les conditions de régularité du processus de naissance et de mort. On se servira du théorème 4, § 3. Le système d'équations (19), § 3, devient

$$\begin{aligned}\lambda g_0&=-\lambda_0 g_0+\lambda_0 g_1,\\ \lambda g_k&=-(\lambda_k+\mu_k)\,g_k+\lambda_k g_{k+1}+\mu_k g_{k-1},\quad k>0.\end{aligned} \tag{13}$$

On voit qu'on peut définir tous les g_k en fonction de g_0. Si g_0 est nul, tous les g_k le seront; si $g_0\neq 0$, le rapport g_k/g_0 est défini de façon univoque à partir de (13). Notons $g_{k+1}-g_k=f_k$. La deuxième expression de (13) donne

$$f_k=\frac{\lambda}{\lambda_k}\,g_k+\frac{\mu_k}{\lambda_k}\,f_{k-1}=$$

$$=\frac{\lambda}{\lambda_k}\,g_k+\frac{\mu_k}{\lambda_k}\left[\frac{\lambda}{\lambda_{k-1}}\,g_{k-1}+\frac{\mu_{k-1}}{\lambda_{k-1}}\left(\frac{\lambda}{\lambda_{k-2}}\,g_{k-2}+\ldots\right)\right]=$$

$$=\frac{\lambda}{\lambda_k}\,g_k+\frac{\mu_k}{\lambda_k}\,\frac{\lambda}{\lambda_{k-1}}\,g_{k-1}+\frac{\mu_k}{\lambda_k}\,\frac{\mu_{k-1}}{\lambda_{k-1}}\,\frac{\lambda}{\lambda_{k-2}}\,g_{k-2}+\ldots$$

$$\ldots+\frac{\mu_k\ldots\mu_2}{\lambda_k\ldots\lambda_2}\,\frac{\lambda}{\lambda_1}\,g_1+\frac{\mu_k\ldots\mu_1}{\lambda_k\ldots\lambda_1}\,f_0. \tag{14}$$

Soit $g_0=1$. Tous les $g_k>0$ et $f_k>0$. Donc g_k croît avec k. Comme $f_0=\frac{\lambda}{\lambda_0}\,g_0$, en remplaçant dans (14) tous les g_k par $g_0=1$, on aura

$$f_k\geqslant\lambda\left[\frac{1}{\lambda_k}+\frac{\mu_k}{\lambda_k\lambda_{k-1}}+\frac{\mu_k\mu_{k-1}}{\lambda_k\lambda_{k-1}\lambda_{k-2}}+\ldots+\frac{\mu_r\ldots\mu_1}{\lambda_k\ldots\lambda_1\lambda_0}\right].$$

La condition

$$\sum_{n=1}^{\infty}\left[\frac{1}{\lambda_n}+\frac{\mu_n}{\lambda_n\lambda_{n+1}}+\ldots+\frac{\mu_n\ldots\mu_1}{\lambda_n\ldots\lambda_1\lambda_0}\right]<\infty \tag{15}$$

est une condition nécessaire d'existence d'une solution bornée du système (13), puisque $\sum_{k=0}^{n} f_k = g_{n+1}-g_0$. Remplaçons maintenant tous les g_t par g_k dans (14). On trouve

$$f_k \leqslant \left[\frac{\lambda}{\lambda_k}+\ldots+\frac{\mu_k\ldots\mu_1\lambda}{\lambda_k\ldots\lambda_1\lambda_0}\right] g_k,$$

$$g_{k+1}\leqslant\left[1+\frac{\lambda}{\lambda_k}+\ldots+\frac{\mu_k\ldots\mu_1\lambda}{\lambda_k\ldots\lambda_1\lambda_0}\right] g_k \leqslant$$

$$\leqslant g_k \exp\left\{\lambda\left[\frac{1}{\lambda_k}+\ldots+\frac{\mu_k\ldots\mu_1}{\lambda_k\ldots\lambda_0}\right]\right\}.$$

Donc, pour $g_0=1$

$$\sup_n g_n \leqslant \exp\left\{\lambda\sum_{k=1}^{\infty}\left[\frac{1}{\mu_k}+\ldots+\frac{\mu_k\ldots\mu_1}{\lambda_k\ldots\lambda_0}\right]\right\}$$

et par suite la relation (15) est une condition suffisante pour que la solution de (13) soit bornée.

Théorème 1. *Pour qu'un processus de naissance et de mort soit régulier il faut et il suffit que*

$$\sum_{n=1}^{\infty}\left[\frac{1}{\lambda_n}+\frac{\mu_n}{\lambda_n\lambda_{n-1}}+\ldots+\frac{\mu_n\ldots\mu_1}{\lambda_n\ldots\lambda_1\lambda_0}\right]=+\infty. \tag{16}$$

Nous allons traiter un cas spécial de processus de naissance et de mort, celui où $\mu_k = 0$, $k = 1, 2, \ldots$ Un tel processus est dit de *reproduction pure* (ou de *croissance*). Ici, de l'état i on ne peut passer qu'à l'état $i+1$. Les réalisations de ce processus sont des fonctions croissantes à valeurs entières dont les sauts sont tous égaux à l'unité. Un tel processus peut servir de modèle mathématique d'enregistrement d'un phénomène se déroulant à des instants aléatoires.

Exemple. Désintégration de la matière. Une matière radioactive (appelée matière-mère) donne naissance à une autre matière radio-active (1-ière fille) d'où naît une 2-ième matière-fille, etc. Fixons un atome de matière-mère. Il se trouve dans l'état initial pendant un intervalle de temps aléatoire, puis se désintègre et se transforme en atome de 1-ière fille, et ainsi de suite. La probabilité de désintégration de l'atome dans l'intervalle (t, s) ne dépend pas

de la durée de vie jusqu'à l'instant t et tout état de l'atome possède une durée moyenne de vie déterminée égale à $l_k = \frac{1}{\lambda_k}$. Un exemple de telles chaînes nous est fourni par la désintégration des isotopes d'uranium et de thorium dont le produit final est un isotope stable de plomb. De ce qui précède il suit que le processus aléatoire décrivant la désintégration de l'atome est markovien.

Les équations de Kolmogorov s'écrivent:

premier système:

$$p'_{ij}(t) = -\lambda_i p_{ij}(t) + \lambda_i p_{i+1,\,j}(t), \quad i \geqslant 0,$$

second système

$$\begin{aligned} p'_{ij}(t) &= -p_{ij}(t)\lambda_j + p_{i,\,j-1}(t)\lambda_{j-1}, \quad j \geqslant i+1, \\ p'_{ii}(t) &= -p_{ii}(t)\lambda_0. \end{aligned} \tag{17}$$

A ces équations il importe d'ajouter les conditions initiales $p_{ij}(0) = \delta_{ij}$. Trouvons la solution de ces équations.

Si $i > j$, conformément à la définition du processus de croissance on posera

$$p_{ij}(t) = 0,$$

ce qui, évidemment, est solution du système (17) vérifiant les conditions initiales. Le système d'équations (17) devient désormais récurrentiel (pour i fixe).

Tirons tout d'abord $p_{ii}(t)$ de l'équation

$$p'_{ii}(t) = \lambda_i p_{ii}(t), \quad p_{ii}(0) = 1,$$

et ensuite successivement les fonctions $p_{i,\,i+1}(t)$, $p_{i,\,i+2}(t)$ pour chacune desquelles (17) est une équation différentielle ordinaire linéaire. On obtient d'abord $p_{ii}(t) = e^{-\lambda_i t}$ et ensuite

$$p_{ij}(t) = \lambda_{j-1} \int_0^t \exp\{-\lambda_j (t-s)\} p_{i,\,j-1}(s)\, ds.$$

La solution analytique du système (17) s'obtient sans peine avec les méthodes classiques du calcul opérationnel.

Considérons à cet effet les transformées de Laplace des fonctions $p_{ij}(t)$:

$$\varphi_{ij}(z) = \int_0^\infty e^{-zt} p_{ij}(t)\, dt.$$

On a

$$\int_0^\infty e^{-zt} p'_{ij}(t)\, dt = z\varphi_{ij}(z) - \delta_{ij}.$$

Les équations (17) deviennent

$$z\varphi_{ij}(z) = -\lambda_j \varphi_{ij}(z) + \lambda_{j-1}\varphi_{i,\,j-1}(z), \quad j > i,$$

$$z\varphi_{ii}(z) - 1 = -\lambda_i \varphi_{ii}(z),$$

d'où

$$\varphi_{ii}(z) = \frac{1}{z+\lambda_i}, \quad \varphi_{ij} = \frac{\lambda_{j-1}}{z+\lambda_j}\varphi_{i,\,j-1}, \ j > i,$$

et

$$\varphi_{ij}(z) = \left(\prod_{k=i}^{j-1} \lambda_k\right) \frac{1}{\psi(z)},$$

où

$$\psi(z) = \prod_{k=i}^{j} (z+\lambda_k).$$

Si les λ_k sont tous distincts, on a

$$\frac{1}{\psi(z)} = \sum_{k=i}^{j} \frac{1}{(z+\lambda_k)\,\psi'(-\lambda_k)}.$$

Cette formule nous permet d'écrire $\varphi_{ij}(z)$ sous la forme

$$\varphi_{ij}(z) = \left(\prod_{k=i}^{j-1} \lambda_k\right) \sum_{k=i}^{j} \frac{1}{(z+\lambda_k)\,\psi'(-\lambda_k)}.$$

Comme $\varphi_{ii}(z) = \frac{1}{z+\lambda_i}$ est la transformée de Laplace de la fonction $e^{-\lambda_i t}$, on trouve

$$p_{ij}(t) = 0 \quad \text{pour } j < i,$$

$$p_{ij}(t) = \left(\prod_{k=i}^{j-1} \lambda_k\right) \sum_{k=i}^{j} \frac{e^{-\lambda_k t}}{\psi'(-\lambda_k)} \quad \text{pour } j > i,$$

$$p_{ii}(t) = e^{-\lambda_i t},$$

où

$$\psi'(-\lambda_k) = \prod_{\substack{r=i \\ r\neq k}}^{j} (\lambda_r - \lambda_k).$$

On a donc obtenu la solution analytique des équations (17).

Le processus est dit à *croissance linéaire* si $\lambda_k = k\lambda$. Dans ce cas

$$p_{ij}(t) = i(i+1)\ldots(j-1)\lambda^{j-i}\times$$
$$\times\sum_{k=i}^{j}\frac{e^{-\lambda kt}}{\lambda^{j-i}(i-k)(i+1-k)\ldots(-1)\cdot 1\ldots(j-k)}=$$
$$=e^{-i\lambda t}C_{j-1}^{j-i}(1-e^{-\lambda t})^{j-i}.$$

La répartition correspondante est dite de Yule-Furry. On remarquera que ($\xi(0) = i$)

$$\mathsf{E}\xi(t) = ie^{\lambda t},\quad \operatorname{Var}\xi(t) = ie^{\lambda t}(e^{\lambda t}-1).$$

La chaîne emboîtée d'un processus de croissance vérifiant

$$\mathsf{P}_{0r}\{x_n(\omega) = n + r\} = 1,$$

le théorème 4, § 2, nous apprend que la condition

$$\sum_{k=0}^{\infty}\frac{1}{\lambda_k} = +\infty$$

est une condition nécessaire et suffisante de régularité du processus. En particulier, le processus à croissance linéaire est régulier.

§ 5. Processus branchus

Les processus branchus constituent une classe importante de processus markoviens à nombre dénombrable d'états.

Supposons qu'on observe un système physique composé d'un nombre fini de particules de même type ou de type distinct. Toute particule peut disparaître ou transmuter indépendamment des autres. Les phénomènes décrits par ce schéma sont assez courants dans la nature et en technique : tel est le cas notamment des orages de rayons cosmiques, de la traversée de la matière par des particules élémentaires, de l'évolution de populations biologiques, de la propagation d'une épidémie, etc.

La définition exacte de tels processus dans le cadre de la théorie des processus markoviens débouche sur la notion de processus branchu.

Soit n le nombre des divers types de particules. L'état du système Σ à l'instant t est caractérisé par un vecteur à valeurs entières $\nu(t) = \{\nu_1(t), \nu_2(t), \ldots, \nu_n(t)\}$, où $\nu_i(t)$ est le nombre de particules du type i, existant à l'instant t. Dans la suite on identifiera l'état du système Σ à l'instant t au vecteur $\nu(t)$. S'agissant du caractère de l'évolution du système Σ, on admettra ce qui suit : quelle que soit la particule considérée à l'instant t, son évolution ne dépend

ni de la date et de la façon dont elle est apparue, ni du caractère de l'évolution des autres particules constituant le système Σ à l'instant t.

Soient α, β, ... des vecteurs n-dimensionnels à coordonnées non négatives entières:

$$\alpha = \{a_1, a_2, \ldots, a_n\}, \quad \beta = \{b_1, b_2, \ldots, b_n\}.$$

Introduisons les probabilités de passage $p_{\alpha\beta}(t_1, t_2)$ du système Σ de l'état α à l'instant t_1 à l'état β à l'instant t_2:

$$p_{\alpha\beta}(t_1, t_2) = \mathsf{P}\{\nu(t_2) = \beta \mid \nu(t_1) = \alpha\}.$$

Soit $\{i\}$ ($i = 1, \ldots, n$) l'état du système Σ composé d'une seule particule du type i.

L'hypothèse faite sur le caractère de l'évolution du système Σ se traduit par la formule

$$p_{\alpha\beta}(t_1, t_2) = \sum_{\sum\limits_{i=1}^{n} \sum\limits_{j=1}^{a_i} \beta^{(ij)} = \beta} \prod_{i=1}^{n} \prod_{j=1}^{a_i} p_{\{i\}\,\beta^{(ij)}}(t_1, t_2), \tag{1}$$

où la sommation est étendue à tous les vecteurs $\beta^{(ij)}$ de coordonnées entières non négatives ($j = 1, \ldots, a_i$, $i = 1, \ldots, n$), qui ont pour résultante le vecteur β. Si $a_i = 0$, on pose

$$\prod_{j=1}^{a_i} p_{\{i\}\,\beta^{(ij)}}(t_1, t_2) = 0.$$

Un *processus branchu* est donc un processus markovien dont l'espace N des états possibles est l'ensemble des vecteurs n-dimensionnels de coordonnées entières non négatives et de probabilités de passage vérifiant (1).

Dans la suite on n'envisage que des processus branchus homogènes, c'est-à-dire des processus tels que

$$p_{\alpha\beta}(t_1, t_2) = p_{\alpha\beta}(t_2 - t_1).$$

Nous allons voir des processus à un type de particules ($n = 1$). Toute particule évoluant indépendamment des autres dans un processus branchu, on peut admettre qu'une seule particule existait à l'instant initial. A la longue soit elle disparaît, soit elle se transforme en k particules de même type (première génération).

Toute particule de la première génération « vit » indépendamment des autres et elle est justiciable des mêmes lois probabilistes que l'initiale. A un certain moment soit elle disparaît, soit elle se transforme en une particule de la deuxième génération, etc. Le processus tout entier est décrit par une fonction aléatoire $\nu(t)$ à valeurs entières, égale au nombre de particules existant à l'instant t ($\nu(0) = 1$).

L'ensemble de tous les états possibles du processus est la suite des nombres naturels 0, 1, 2, . . ., l'état 0 étant absorbant: si $\nu(t_0) = 0$, alors $\nu(t) = 0$ $(t > t_0)$. Le processus est dit *dégénérant* si $\nu(t) = 0$ presque sûrement. Dans les autres cas $\nu(t) = 0$ avec une certaine probabilité au bout d'un intervalle de temps fini. Cette probabilité s'appelle *probabilité de dégénération* du processus. Il est possible que $\nu(t)$ croisse indéfiniment avec t. Dans le cas d'une réaction nucléaire cette situation peut être interprétée comme une explosion. Donc, en théorie des processus branchus on s'intéressera aux problèmes suivants: quelle est la probabilité de dégénération d'un processus branchu? Quel est le comportement asymptotique de $\nu(t)$?

Soit $p_{ij}(t)$ probabilité conditionnelle que le système comporte j particules à l'instant $t + \tau$ sachant qu'à l'instant t il en comptait i. La méthode des fonctions génératrices se prête bien à l'étude des problèmes de la théorie des processus branchus. Soient $f_i(z, t)$ fonctions génératrices des répartitions $\{p_{ij}(t)\}$, $j = 0, 1, \ldots$:

$$f_i(z, t) = \sum_{k=0}^{\infty} z^k p_{ik}(t), \quad |z| \leqslant 1.$$

Les probabilités $p_{ij}(t)$ (i fixe) sont associées à la répartition de la somme de i variables aléatoires indépendantes, de même répartition et de fonction génératrice $f_i(z, t)$. Donc

$$f_i(z, t) = [f_1(z, t)]^i. \tag{2}$$

Pour déterminer la fonction $f_1(z, t)$ on pourra se servir de n'importe quel système d'équations de Kolmogorov (§ 3).

La théorie générale affirme l'existence des dérivées

$$\lim_{t\to 0} \frac{p_{1j}(t)}{t} = b_j, \quad j > 1, \quad \lim_{t\to 0} \frac{1-p_{11}(t)}{t} = b_1.$$

Supposons que $b_1 = b_0 + \sum_{j=2}^{\infty} b_j < \infty$. Donc a lieu le premier système d'équations de Kolmogorov:

$$\frac{dp_{1j}(t)}{dt} = -b_1 p_{1j}(t) + \sum_{\substack{k=0\\k\neq 1}}^{\infty} b_k p_{kj}(t). \tag{3}$$

En multipliant les deux membres de (3) par z^j et en sommant sur j de 0 à ∞, on trouve

$$\frac{\partial f_1(z, t)}{\partial t} = -b_1 f_1(z, t) + \sum_{\substack{k=0\\k\neq 1}}^{\infty} b_k f_k(z, t) \quad (|z| \leqslant 1)$$

ou en vertu de (2)

$$\frac{\partial f(z, t)}{\partial t} = -b_1 f(z, t) + \sum_{\substack{k=0 \\ k \neq 1}}^{\infty} b_k f^k(z, t),$$

où $f(z, t) = f_1(z, t)$. En définitive on obtient l'équation différentielle non linéaire suivante:

$$\frac{\partial f}{\partial t} = u(f), \tag{4}$$

où

$$u(z) = b_0 - b_1 z + \sum_{k=2}^{\infty} b_k z^k \quad (|z| \leqslant 1). \tag{5}$$

A l'équation (4) il faut ajouter la condition initiale

$$f(z, 0) = z. \tag{6}$$

La solution de l'équation (4) qui vérifie la condition (6) s'écrit

$$\varphi(f) - \varphi(z) = t, \quad \varphi(z) = \int \frac{dz}{u(z)}. \tag{7}$$

Supposons que sont satisfaites les conditions d'existence du second système d'équations différentielles de Kolmogorov. A noter que la définition des processus branchus nous conduit aux formules suivantes:

$$p_{kk}(t) = (1 - b_1 t)^k + o(t) = 1 - kb_1(t) + o(t),$$
$$p_{k, k-1}(t) = k(1 - b_1 t)^{k-1} b_0 t + o(t) = kb_0 t + o(t),$$
$$p_{k, k-j}(t) = o(t), \quad j \geqslant 2,$$
$$p_{k, k+j}(t) = k(1 - b_1 t) b_{j+1} t + o(t) = kb_{j+1} t + o(t), \quad j \geqslant 1,$$

d'où il suit (dans les notations du § 3)

$$a_{jj} = jb_1, \; a_{j, j-1} = jb_0, \; a_{j, j-k} = 0, \; k \geqslant 2;$$
$$a_{j, j+k} = jb_{k+1}, \; k \geqslant 1.$$

Donc, pour les fonctions $p_{1j}(t)$, le système d'équations (15), § 3, s'écrit

$$\frac{\partial p_{1j}(t)}{\partial t} = (j+1) b_0 p_{1, j+1}(t) - jb_1 p_{1j}(t) + $$
$$+ (j-1) b_2 p_{1, j-1}(t) + \ldots + b_j p_{11}(t). \tag{8}$$

En multipliant (8) par z^j et en sommant sur j de 0 à ∞, on obtient une nouvelle équation pour la fonction génératrice $f(z, t)$:

$$\frac{\partial f(z, t)}{\partial t} = u(z) \frac{\partial f(z, t)}{\partial z} \quad (|z| < 1), \tag{9}$$

où $u(z)$ est définie par (5). A cette équation il convient d'ajouter la condition initiale (6). La solution des équations (9) vérifiant la condition (6) est

$$f(z, t) = \psi\left(t + \int_0^z \frac{dz}{u(z)}\right), \tag{10}$$

où $\psi(t)$ est l'inverse de $t = \varphi(z) = \int_0^z \frac{dz}{u(z)}$. Cette solution est confondue avec (7).

Exemple. Posons

$$u(z) = p - (p+q) z + qz^2.$$

Dans ce cas, au bout d'un intervalle de temps t, une particule disparaît avec la probabilité $pt + o(t)$, ou bien se transforme en deux particules avec la probabilité $qt + o(t)$, ou bien encore se conserve avec la probabilité $1 - (p+q)t + o(t)$. La probabilité de transmutation en plus de deux particules vaut $o(t)$. On a

$$\varphi(z) = \int_0^z \frac{dz}{u(z)} = \int_0^z \frac{dz}{p-(p+q)z+qz^2} = \frac{1}{q-p} \ln \frac{z-1}{z-\frac{p}{q}} \quad (p \neq q).$$

Pour la fonction inverse $z = \psi(t)$ on obtient l'expression

$$z = \psi(t) = \frac{1-\beta e^{qt(1-\beta)}}{1-e^{qt(1-\beta)}}, \quad \beta = \frac{p}{q},$$

d'où

$$f(z, t) = \psi(t + \varphi(z)) = \frac{z-\beta-\beta(z-1)e^{qt(1-\beta)}}{z-\beta-(z-1)e^{qt(1-\beta)}}.$$

En développant $f(z, t)$ en série sur les puissances de z:

$$f(z, t) = \frac{1-e^{t(q-p)}}{1-\frac{1}{\beta}e^{t(q-p)}} + \sum_{n=1}^{\infty} \frac{(1-\beta)^2(1-e^{t(q-p)})^{n-1}e^{t(q-p)}}{[\beta-e^{t(q-p)}]^{n+1}} z^n,$$

on est conduit aux formules:

$$p_{1\,0}(t) = \frac{1-e^{t(q-p)}}{1-\frac{1}{\beta}e^{t(q-p)}}, \tag{11}$$

$$p_{1\,n}(t) = (1-\beta)^2 \frac{(1-e^{t(q-p)})^{n-1}e^{t(q-p)}}{[\beta-e^{t(q-p)}]^{n+1}}, \quad n = 1, 2, \ldots \tag{12}$$

Traitons le cas $p=q$ qui a été exclu précédemment. On a

$$\varphi(z)=\int_0^z \frac{dz}{p(1-z)^2}=\frac{1}{p}\frac{z}{1-z}, \quad z=\psi(t)=1-\frac{1}{1+pt},$$

d'où

$$f(z, t)=z(t+\varphi(z))=1-\frac{1-z}{1+pt(1-z)}=\frac{pt}{1+pt}+\sum_{n=1}^{\infty}\frac{(pt)^{n-1}}{(1+pt)^{n+1}}z^n.$$

Donc

$$p_{10}(t)=\frac{pt}{1+pt}, \tag{13}$$

$$p_{1n}(t)=\frac{(pt)^{n-1}}{(1+pt)^{n+1}}, \quad n=1, 2, \ldots \tag{14}$$

Les formules (11) à (14) entraînent aussitôt les relations asymptotiques suivantes pour $t\to\infty$: si $q \leqslant p$, alors

$$p_{10}(t)\to 1, \; p_{1n}(t)\to 0 \text{ pour } n\geqslant 1, t\to\infty;$$

si $q>p$, alors

$$p_{10}(t)\to\beta\,(\beta<1), \; p_{1n}(t)\to 0 \text{ pour } n\geqslant 1, \; t\to\infty.$$

Dans le premier cas ($q\leqslant p$) le processus branchu dégénère presque sûrement, c'est-à-dire toutes les particules finissent par disparaître. Dans le second cas ($q>p$) la probabilité de dégénération du processus est $\beta=\frac{p}{q}<1$; si les particules ne disparaissent pas, leur nombre croît indéfiniment avec t. En effet,

$$\mathsf{P}\{\nu(t)>N \mid \nu(t)>0\}=1-\frac{1}{p_{10}(t)}\sum_{k=1}^{N}p_{1k}(t)\to 1, \; \forall N.$$

Etudions le comportement asymptotique pour $t\to\infty$ d'un processus branchu dans le cas général. On aura besoin des moments de $\nu(t)$.

Cependant, avec les fonctions génératrices il est préférable d'utiliser les moments factoriels. Posons

$$m_k(t)=\mathsf{E}\,[\nu(t)(\nu(t)-1)\ldots(\nu(t)-k+1)].$$

On établit sans peine les équations différentielles linéaires vérifiées par les moments factoriels $m_k(t)$. Supposons que

$$\sum_{k=1}^{\infty}kb_k<\infty.$$

Alors pour $|z| < 1$ une dérivation par rapport à z de l'équation (4) donne

$$\frac{\partial m_1(z, t)}{\partial t} = u'(f)\, m_1(z, t), \tag{15}$$

où

$$m_1(z, t) = \frac{\partial}{\partial z} f(z, t) = \mathsf{E}\nu(t)\, z^{\nu(t)},$$

$$u'(z) = -b_1 + \sum_{k=2}^{\infty} k b_k z^{k-1}.$$

De (15) et de $m_1(z, 0) = 1$ il résulte

$$m_1(z, t) = e^{\int_0^t u'(f)\, dt} \qquad (|z| < 1).$$

Pour $z \to 1$ on a $u'(z) \to -b_1 + \sum_{k=2}^{\infty} k b_k = m_1 < \infty$ et $f(z, t) \uparrow 1$ de façon monotone et partant uniformément en t. Par suite

$$\lim_{z \uparrow 1} m_1(z, t) = e^{m_1 t}.$$

D'autre part, le théorème de Lebesgue nous dit que pour $z \uparrow 1$

$$\lim_{z \uparrow 1} m_1(z, t) = \lim_{z \uparrow 1} \mathsf{E}\nu(t)\, z^{\nu(t)} = \mathsf{E}\nu(t) = \sum_{k=1}^{\infty} k p_k(t).$$

Donc

$$m_1(t) = e^{m_1(t)}. \tag{16}$$

Ce résultat se généralise immédiatement aux moments factoriels d'ordres supérieurs.

LEMME 1. *Si*

$$m_r = \sum_{n=1}^{\infty} n(n-1) \ldots (n-r+1)\, b_n < \infty, \quad r = 1, \ldots, k, \tag{17}$$

les moments factoriels $m_r(t)$, $r = 1, 2, \ldots, k$, de $\nu(t)$ sont finis et solutions des équations différentielles linéaires du premier ordre à coefficients constants.

Pour le prouver dérivons l'équation (4) successivement par rapport à z $(|z| < 1)$. En posant

$$m_k(z, t) = \frac{\partial^k f(z, t)}{\partial z^k} = \mathsf{E}\left(\nu(t)(\nu(t)-1) \ldots (\nu(t)-k+1)\, z^{\nu(t)-k}\right),$$

on trouve

$$\begin{aligned} \frac{\partial m_2(z, t)}{\partial t} &= u'(f)\, m_2(z, t) + u''(f)\, m_1^2(z, t), \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \frac{\partial m_k(z, t)}{\partial t} &= u'(f)\, m_k(z, t) + F_k(z, t), \end{aligned} \tag{18}$$

où $F_k(z, t)$ est un polynôme en $m_1(z, t), \ldots, m_{k-1}(z, t)$ dont les coefficients dépendent linéairement de $u''(f), \ldots, u^{(k)}(f)$. A ces équations il faut ajouter les conditions initiales

$$m_k(z, 0) = 0, \quad k = 2, 3, \ldots$$

La solution de l'équation (18) est de la forme

$$m_k(z, t) = e^{\int_0^t u'(f)\,dt} \int_0^t F_k(z, \tau)\, e^{-\int_0^\tau u'(f)\,d\theta}\, d\tau.$$

En raisonnant par récurrence et en se servant des mêmes considérations que pour le cas $k = 1$, on obtient

$$m_k(t) = \lim_{z\to 1} m_k(z, t) = e^{m_1 t} \int_0^t F_k(1, \tau)\, e^{-m_1\tau}\, d\tau\,;$$

on voit que $m_k(t)$ est solution de l'équation (18) avec $z = 1$. ∎

En particulier,

$$m_2(t) = \frac{m_2}{m_1}(e^{m_1 t} - 1)\, e^{m_1 t} \quad (m_1 \neq 0), \tag{19}$$

$$m_2(t) = m_2 t \quad (m_1 = 0). \tag{20}$$

La fonction $u(z)$ (cf. (5)) est importante dans l'étude du comportement asymptotique du processus branchu. Considérons cette fonction pour les valeurs réelles de z. A noter que

$$u(0) = b_0 \geqslant 0, \quad u(1) = b_0 - b_1 + \sum_{k=2}^{\infty} b_k = 0,$$
$$u''(z) > 0 \text{ pour } z > 0\,;$$

on admettra que les b_k $(k \geqslant 2)$ ne sont pas tous nuls. Donc $u''(z)$ est convexe vers le bas pour $z > 0$ et par suite possède au plus un zéro sur l'intervalle $]0, 1[$. Passons à la définition de la probabilité α de dégénération du processus branchu $\nu(t)$. Les événements $\{\nu(t) = 0\}$ formant une classe monotone croissante, on a

$$\alpha = \mathsf{P}\{\lim_{t\to\infty} \nu(t) = 0\} = \lim_{t\to\infty} \mathsf{P}\{\nu(t) = 0\} = \lim_{t\to\infty} p_{10}(t).$$

THÉORÈME 1. *La probabilité de dégénération d'un processus branchu est confondue avec la plus petite racine non négative de l'équation* $u(x) = 0$. *Si*

$$u'(1) = m_1 = -b_1 + \sum_{k=2}^{\infty} k b_k < \infty,$$

alors

$$\alpha \begin{cases} =1 & \textit{pour } u'(1)\leqslant 0; \\ <1 & \textit{pour } u'(1)>0. \end{cases}$$

D é m o n s t r a t i o n. Puisque $p_{1\,0}(t) = f(0, t)$, de (4) il suit

$$\frac{dp_{10}(t)}{dt} = u(p_{10}(t)), \quad p_{10}(0) = 0. \tag{21}$$

Si $b_0 = 0$, $p_{1\,0}(t) \equiv 0$ est solution de l'équation (21) et le théorème 1 est trivial. Soit $b_0 > 0$. A noter que si x_0 est la plus petite racine positive de l'équation $u(x) = 0$, alors $p_{1\,0}(t_0) < x_0$, $\forall t > 0$. En effet, si $p_{1\,0}(t_0) = x_0$, $t_0 > 0$, en vertu de l'unicité de la solution de l'équation (21) on aurait $p_{1\,0}(t) \equiv x_0$, ce qui est impossible.

Comme d'autre part existe $\alpha = \lim\limits_{t\to\infty} p_{10}(t) \leqslant 1$, de (21) suit l'existence de $\lim\limits_{t\to\infty} p'_{10}(t) = u(\alpha)$. D'où il résulte que $u(\alpha) = 0$ sinon $p_{10}(t) = \int\limits_{t_0}^{t} p'_{10}(t)\,dt + p_{10}(t_0)$ croîtrait indéfiniment. On a donc montré que $\alpha = x_0$. Si $x_0 < 1$, la fonction $u(x)$ est croissante au point $x = 1$ et la dérivée $u'(1)$, si elle existe, est >0. Si $x_0 = 1$, alors $u'(1) \leqslant 0$. ■

Etudions maintenant le comportement asymptotique de la probabilité $p_{1\,0}(t)$ pour $t \to \infty$ pour les processus dégénérants ($\alpha = 1$).

THÉORÈME 2. *Si* $m_1 = u'(1) \leqslant 0$, $m_2 = u''(1) < \infty$, *alors*

$$1 - p_{10}(t) \sim \begin{cases} Ke^{m_1 t} & \textit{pour } m_1 < 0 \textit{ et un } K > 0; \\ \dfrac{2}{m_2 t} & \textit{pour } m_1 = 0. \end{cases}$$

D é m o n s t r a t i o n. Posons

$$q(t) = 1 - p_{1\,0}(t).$$

La fonction $q(t)$ est solution de l'équation

$$\frac{dq}{dt} = -u(1 - q(t)), \quad q(0) = 1.$$

La formule des accroissements finis nous montre que

$$\frac{dq}{dt} = -u(1) + q(t)\,u'(\xi) = q(t)\,u'(\xi),$$

où ξ est compris entre $p_{1\,0}(t)$ et 1. Comme $u'(x)$ est une fonction monotone croissante et que $\xi \to 1$ pour $t \to \infty$, on a $u'(\xi) = u'(1) - \varepsilon(t)$, où $\varepsilon(t) > 0$ et $\lim\limits_{t\to\infty} \varepsilon(t) = 0$. Par suite

$$\frac{dq}{dt} = q(t)\,(m_1 - \varepsilon(t)),$$

d'où

$$q(t)=e^{m_1 t-\int\limits_0^t \varepsilon(\tau)\,d\tau}.$$

On remarquera que $(\xi<\zeta<1)$

$$0<\varepsilon(t)=u'(1)-u'(\xi)=u''(\zeta)(1-\xi)\leqslant$$
$$\leqslant u''(1)(1-p_{1\,0}(t))\leqslant m_2 e^{m_1 t},$$

donc l'intégrale $\int\limits_0^\infty \varepsilon(t)\,dt$ est finie. D'où il résulte que pour $m_1<0$

$$q(t)\sim Ke^{m_1 t},\quad \text{où}\quad K=\exp\left(-\int\limits_0^\infty \varepsilon(t)\,dt\right).$$

Traitons le cas $m_1=0$. On a

$$\frac{dq}{dt}=-u(1-q(t))=-u(1)+q(t)u'(1)-\frac{q^2(t)}{2}u''(\xi_1),$$

où $\xi_1\in\,]p_{1\,0}(t),\ 1[$. Comme $u''(\xi_1)\to u''(1)$ pour $t\to\infty$, il vient

$$\frac{dq}{dt}=-\frac{q^2(t)}{2}(m_2+\varepsilon(t)),$$

où $\varepsilon(t)\to 0$ pour $t\to\infty$. D'où il suit

$$q(t)=\frac{2}{m_2 t+\int\limits_0^t \varepsilon(\tau)\,d\tau+2}=\frac{2}{m_2 t}+o\left(\frac{1}{t}\right).\ \blacksquare$$

Complétons le théorème 2 des résultats relatifs au comportement asymptotique des probabilités $p_{1\,k}(t)$ pour des processus dégénérants.

$\lim\limits_{t\to\infty} p_{1\,n}(t)=0$ $(n>0)$ implique $\lim\limits_{t\to\infty} f(z,t)=1$. Posons

$$q(z,t)=1-f(z,t).$$

Pour $z=0$ on a $q(0,t)=1-f(0,t)=1-p_{1\,0}(t)=q(t)\sim$ $\sim Ke^{m_1 t}$. Admettons que telle sera la vitesse de décroissance de la fonction $q(z,t)$ pour $z\neq 0$. Posons alors

$$\varphi(z,t)=\frac{q(z,t)}{q(t)}=\frac{1-f(z,t)}{q(t)}. \qquad (22)$$

On remarquera que la fonction

$$f^*(z,t)=1-\varphi(z,t)=\sum_{n=1}^{\infty}\frac{p_{1n}(t)}{q(t)}z^n \qquad (23)$$

peut être considérée comme la fonction génératrice de la répartition conditionnelle du nombre $\nu(t)$ de particules sachant que $\nu(t) \neq 0$.

THÉORÈME 3. *Si* $m_1 = u'(1) < 0$, $m_2 = u''(1) < \infty$, *alors pour* $t \to \infty$ *la répartition conditionnelle du nombre* $\nu(t)$ *de particules sachant que* $\nu(t) > 0$ *tend vers une limite dont la fonction génératrice* $f^*(z)$ *est égale à*

$$f^*(z) = 1 - e^{m_1 \int_0^z \frac{dz}{u(z)}}. \tag{24}$$

D é m o n s t r a t i o n. Soit la fonction $\varphi(z, t)$. De (4) il résulte que $\varphi(z, t)$ est solution de l'équation

$$\frac{\partial \varphi}{\partial t} = -\frac{1}{q(t)} u(1 - q(t)\varphi) + \frac{\varphi}{q(t)} u(1 - q(t)).$$

Le développement du second membre en série de Taylor s'écrit:

$$\frac{\partial \varphi}{\partial t} = -\frac{1}{q(t)} \left[u(1) - q(t)\varphi u'(1) + \frac{(q(t)\varphi)^2}{2}(u''(1) + \varepsilon_1) \right] +$$
$$+ \frac{\varphi}{q(t)} \left[u(1) - q(t) u'(1) + \frac{q^2(t)}{2}(u''(1) + \varepsilon_2) \right],$$

où $\varepsilon_1 = u''(\xi_1) - u''(1)$, $\varepsilon_2 = u''(\xi_2) - u''(1)$, ξ_1 (resp. ξ_2) un nombre compris entre $f(z, t)$ et 1 (resp. entre $f(0, t)$ et 1). Les fonctions ε_i ($i = 1, 2$) tendent uniformément vers zéro dans tout domaine $|z| \leqslant \rho < 1$ pour $t \to \infty$. L'égalité précédente s'écrit encore

$$\frac{\partial \varphi}{\partial t} = -\frac{q(t)\varphi^2}{2}(m_2 + \varepsilon_1) + \frac{q(t)\varphi}{2}(m_2 + \varepsilon_2). \tag{25}$$

A partir d'un t assez grand on aura

$$\frac{\partial \varphi}{\partial t} < \frac{q(t)\varphi}{2}\left(m_2 + \frac{m_2}{2}\right) = \frac{3}{4} m_2 q(t)\varphi,$$

d'où

$$\varphi(z, t) \leqslant \varphi(z, t_0)\, e^{\frac{3}{4} m_2 \int_{t_0}^t q(\tau)\, d\tau}.$$

La convergence de l'intégrale $\int_{t_0}^\infty q(\tau)\, d\tau$ indique que la fonction $\varphi(z, t)$ reste bornée pour $t \to \infty$. Donc on peut écrire l'équation (25) sous la forme

$$\frac{\partial \varphi}{\partial t} = \frac{q(t)\varphi}{2}[m_2(1 - \varphi) + \varepsilon], \qquad \varphi(z, 0) = 1 - z,$$

où $\varepsilon = \varepsilon_2 - \varphi\varepsilon_1 \to 0$ pour $t \to \infty$. En mettant la solution de la dernière équation sous la forme

$$\varphi(z, t) = (1 - z)\, e^{\frac{1}{2}\int_0^t q(\tau)[m_2(1 - \varphi(z, \tau)) + \varepsilon]\, d\tau}$$

on s'assure de l'existence de la limite

$$\lim_{t\to\infty} \varphi(z, t) = K(z).$$

De (25) il suit que $\lim\limits_{t\to\infty} \frac{\partial\varphi}{\partial t} = 0$. Comme $\varphi(z, t)$ est une fonction analytique dans le disque $|z| < 1$ et que toutes les limites utilisées ont lieu uniformément dans tout disque $|z| \leqslant \rho < 1$, la fonction $K(z)$ est aussi analytique dans $|z| < 1$ et

$$\lim_{t\to\infty} \frac{\partial\varphi(z, t)}{\partial z} = \frac{dK(z)}{dz}$$

uniformément dans tout disque $|z| \leqslant \rho < 1$. Pour déterminer la fonction $K(z)$ on se sert de l'équation (9). En y portant $f(z, t) = 1 - q(t)\varphi(z, t)$, on trouve

$$-q'(t)\varphi(z, t) - q(t)\frac{\partial\varphi(z, t)}{\partial t} = -u(z)q(t)\frac{\partial\varphi(z, t)}{\partial z}.$$

En divisant par $q(t)$ pour $t\to\infty$, et puisque $q'(t)/q(t) \to m_1$ (cf. démonstration du théorème 2), on obtient

$$m_1 K(z) = u(z)\frac{dK(z)}{dz},$$

avec $K(0) = \lim\limits_{t\to\infty} \varphi(0, t) = 1$. Donc

$$K(z) = e^{m_1 \int_0^z \frac{dz}{u(z)}},$$

$$1 - f(z, t) \sim q(t)K(z) = e^{m_1\left(t + \int_0^z \frac{dz}{u(z)}\right)}. \blacksquare$$

Nous allons établir l'expression du nombre moyen de particules à l'instant t sachant que le processus n'a encore pas dégénéré. On a

$$m^*(t) = \mathsf{E}\{\nu(t) \mid \nu(t) > 0\} = \frac{e^{m_1 t}}{q(t)}, \tag{26}$$

d'où, en vertu du théorème 2, l'on déduit les relations asymptotiques pour $t\to\infty$:

$$m^*(t) \sim \begin{cases} 1/K & \text{pour } m_1 < 0, \\ m_2 t/2 & \text{pour } m_1 = 0, \\ e^{m_1 t}/(1-\alpha) & \text{pour } m_1 > 0. \end{cases}$$

Pour $m_1 \geqslant 0$ le nombre $\nu(t)$ de particules sachant que $\nu(t) > 0$ croît indéfiniment. Posons

$$\nu^*(t) = \frac{\nu(t)}{m^*(t)}.$$

On a alors $\mathsf{E}\{\nu^*(t) \mid \nu^*(t) > 0\} = 1$. Etudions le comportement de la limite de $\nu^*(t)$ pour $t \to \infty$.

On a intérêt à passer aux fonctions caractéristiques puisque tout laisse à croire que, si elle existe, la répartition limite de $\nu^*(t)$ sachant que $\nu(t) > 0$ sera répartition continue sur la section $[0, \infty[$. Pour la fonction caractéristique $g(\lambda, t)$ de la variable aléatoire $\nu^*(t)$ sachant que $\nu(t) > 0$ on a:

$$g(\lambda, t) = \sum_{n=1}^{\infty} e^{\frac{i\lambda n}{m^*(t)}} \frac{p_{1n}(t)}{q(t)} = \frac{f\left(e^{\frac{i\lambda}{m^*(t)}}, t\right) - f(0, t)}{q(t)},$$

ou

$$g(\lambda, t) = 1 - \frac{1 - f\left(e^{\frac{i\lambda}{m^*(t)}}, t\right)}{q(t)}. \tag{27}$$

Traitons le cas $m_1 = 0$.

THÉORÈME 4. *Si* $m_1 = 0$ *et* $m_2 < \infty$, *alors*

$$\lim_{t\to\infty} \mathsf{P}\left\{\frac{2\nu(t)}{m_2(t)} < x \mid \nu(t) > 0\right\} = 1 - e^{-x}. \tag{28}$$

D é m o n s t r a t i o n. En posant

$$\psi(z, t) = 1 - f(z, t),$$

on déduit de l'équation (4)

$$\frac{\partial\psi}{\partial t} = -u(1-\psi) = -\frac{\psi^2}{2}[u''(1) + \varepsilon(t)], \quad \psi(z, 0) = 1 - z.$$

Le processus étant dégénérant pour $m_1 = 0$, on a $\psi(z, t) \to 0$ pour $t \to \infty$ uniformément dans le disque $|z| \leqslant 1$. D'où il résulte que $\varepsilon(t) \to 0$ pour $t \to \infty$ uniformément en z, $|z| \leqslant t$. Une intégration de la dernière équation donne

$$\frac{1}{\psi(z, t)} - \frac{1}{1-z} = \frac{1}{2}\left[m_2 t + \int_0^t \varepsilon(\tau)\, d\tau\right],$$

d'où

$$\lim_{t\to\infty} \frac{q(t)}{\psi\left(e^{\frac{i\lambda}{m^*(t)}}, 1\right)} =$$

$$= \lim_{t\to\infty} \left\{ \frac{q(t)}{\frac{i\lambda}{1-e^{m^*(t)}}} + \frac{m_2 t q(t)}{2} + q(t) \int_0^t \varepsilon(\tau)\, d\tau \right\} = -\frac{1}{i\lambda} + 1$$

et

$$g(\lambda) = \lim_{t\to\infty} g(\lambda, t) = \frac{1}{1-i\lambda}.$$

La fonction $g(\lambda)$ est fonction caractéristique de la répartition $F(x) = 1 - e^{-x}$ pour $x > 0$, $F(x) = 0$ pour $x < 0$. ■

Pour $m_1 > 0$, $q(t)$ tend vers $1 - \alpha \neq 0$. La normalisation à l'aide de la fonction $g(t)$ ou le passage aux espérances mathématiques conditionnelles sachant que $\nu(t) > 0$ ne présentent donc pas grand intérêt. Prouvons le théorème suivant:

THÉORÈME 5. *Si $m_1 > 0$ et $m_2 < \infty$, la quantité $\nu(t)\, e^{-m_1 t}$ converge en moyenne quadratique pour $t \to \infty$ vers $\eta = \text{l.i.m.}\ \nu(t) e^{-m_1 t}$, dont la fonction caractéristique $g(\lambda)$ est solution de l'équation fonctionnelle*

$$(1 - g(\lambda)) \exp\left\{ - \int_1^{g(\lambda)} \frac{u(v) - m_1(v-1)}{u(v)(v-1)}\, dv \right\} = -i\lambda. \qquad (29)$$

On se servira du critère de Cauchy pour prouver la convergence en moyenne quadratique de $\tilde{\nu}(t) = \nu(t)\, e^{-m_1 t}$. Soit $t < t'$. On a

$$\mathsf{E}(\tilde{\nu}(t) - \tilde{\nu}(t'))^2 = \mathsf{E}\tilde{\nu}(t)^2 + \mathsf{E}\tilde{\nu}(t')^2 - 2\mathsf{E}(\tilde{\nu}(t)\tilde{\nu}(t')).$$

De (16) et (19) il suit

$$\mathsf{E}\tilde{\nu}(t)^2 \sim m_2/m_1 \quad \text{pour } t \to \infty.$$

La définition et l'homogénéité du processus branchu nous conduisent à

$$\mathsf{E}(\nu(t)\nu(t')) = \mathsf{E}\nu(t)\, \mathsf{E}\{\nu(t') \mid \nu(t)\} = \mathsf{E}\nu^2(t)\, \mathsf{E}\nu(t'-t),$$

d'où, en vertu des mêmes formules (16) et (19), il résulte

$$\mathsf{E}(\tilde{\nu}(t)\tilde{\nu}(t')) \sim m_2/m_1.$$

Donc $\mathsf{E}(\tilde{\nu}(t) - \tilde{\nu}(t'))^2 \to 0$ pour $t, t' \to \infty$ et existe $\eta = \underset{t\to\infty}{\text{l.i.m.}}\ \nu(t)$.

En mettant l'équation (4) sous la forme

$$\frac{df}{f-1} - \frac{u(f) - m_1(f-1)}{u(f)(f-1)}\, df = m_1\, dt$$

et en intégrant par rapport à t de 0 à t, on trouve

$$\ln(1-f) - \int_0^t \frac{u(v) - m_1(v-1)}{u(v)(v-1)}\, dv = m_1 t + \ln(1-z).$$

Si l'on pose $z = e^{\frac{i\lambda}{m^*(t)}}$ et que l'on fasse tendre t vers ∞, on est conduit à l'égalité

$$\ln(1-g) - \int_0^g \frac{u(v) - m_1(v-1)}{u(v)(v-1)} dv = \ln(-i\lambda),$$

d'où résulte (29). ■

La théorie des processus branchus avec des particules de plusieurs types est analogue mais plus compliquée. Nous allons nous appesantir sur les relations fondamentales de cette théorie. Comme dans le cas de particules d'un seul type, il est plus commode de se servir de la méthode des fonctions génératrices.

Soit M ensemble des états possibles du processus, c'est-à-dire l'ensemble de tous les vecteurs $\alpha = (a_1, a_2, \ldots, a_n)$ de composantes entières non négatives (on désignera les vecteurs n-dimensionnels par des lettres grecques α, β, σ, ..., et leurs composantes par les lettres latines correspondantes). Définissons les fonctions génératrices $F_i(t, \sigma) = F_i(t, s_1, s_2, \ldots, s_n)$ des probabilités de passage $p_{\{i\}\beta}(t)$:

$$F_i(t, \sigma) = F_i(t, s_1, \ldots, s_n) = \sum_{\beta \in M} p_{\{i\}\beta}(t)\, s_1^{b_1} s_2^{b_2} \ldots s_n^{b_n} \qquad (30)$$

$$(\sigma = \{s_1, s_2, \ldots, s_n\}, \quad \beta = \{b_1, b_2, \ldots, b_n\}).$$

On rappelle que $\{i\}$ désigne le vecteur $\{i\} = \{\delta_{i1}, \delta_{i2}, \ldots, \delta_{in}\}$. Les fonctions $F_i(t, \sigma)$ sont des fonctions analytiques de $s_1, s_2, \ldots, s_n$ dans le domaine $|s_i| < 1$ $(i = 1, \ldots, n)$, et de plus

$$\begin{gathered} |F_i(t, \sigma)| \leqslant 1 \quad \text{pour} \quad |s_i| \leqslant 1, \\ F_i(t, 1, \ldots, 1) = 1, \quad F_i(0, \sigma) = s_i. \end{gathered} \qquad (31)$$

Si l'on introduit la fonction vectorielle n-dimensionnelle

$$\Phi(t, \sigma) = \{F_1(t, \sigma), \ldots, F_n(t, \sigma)\},$$

de (31) il résulte

$$\Phi(0, \sigma) = \sigma. \qquad (32)$$

Trouvons maintenant en termes de fonctions génératrices l'équivalent de la formule de Chapman-Kolmogorov pour les processus branchus. On a

$$p_{\{i\}\beta}(t+\tau) = \sum_{\alpha \in M} p_{\{i\}\alpha}(t)\, p_{\alpha\beta}(\tau), \quad t > 0, \ \tau > 0.$$

En remplaçant $p_{\alpha\beta}(\tau)$ par son expression (1), on obtient

$$p_{\{i\}\beta}(t+\tau) = \sum_{\alpha \in M} p_{\{i\}\alpha}(t) \sum_{\sum_k \sum_j \beta^{(k,j)} = \beta} \prod_{k=1}^{n} \prod_{j=1}^{\alpha_i} p_{\{k\}\beta^{(k,j)}}(\tau).$$

En multipliant les deux membres par $s_1^{b_1}, \ldots, s_n^{b_n}$ et en sommant sur tous les β, on trouve

$$F_i(t+\tau, \sigma) =$$

$$= \sum_{\alpha \in M} p_{\{i\}\alpha}(t) \sum_{\beta \in M} \sum_{\sum\limits_k \sum\limits_j \beta^{(k,j)} = \beta} \prod_{k=1}^{n} \prod_{j=1}^{a_k} p_{\{k\}\beta^{(k,j)}}(\tau) s_1^{b_1^{(k,j)}} \ldots s_n^{b_n^{(k,j)}} =$$

$$= \sum_{\alpha \in M} p_{\{i\}\alpha}(t) \prod_{k=1}^{n} \prod_{j=1}^{a_k} \Big(\sum_{\beta^{(k,j)} \in M} p_{\{k\}\beta^{(k,j)}}(\tau) s_1^{b_1^{(k,j)}} \ldots s_n^{b_n^{(k,j)}} \Big) =$$

$$= \sum_{\alpha \in M} p_{\{i\}\alpha}(t) \prod_{k=1}^{n} F_k^{a_k}(\tau, \sigma),$$

d'où il suit

$$F_i(t+\tau, \sigma) = F_i(t, F_1(\tau, \sigma), \ldots, F_n(\tau, \sigma)), \; i = 1, \ldots, s,$$

ou

$$\Phi(t+\tau, \sigma) = \Phi(t, \Phi(\tau, \sigma)). \tag{33}$$

Théorème 6. *Le système de fonctions génératrices d'un processus branchu est solution du système d'équations fonctionnelles* (33) *avec la condition initiale* (32).

Etablissons pour les fonctions génératrices des équations différentielles correspondant au premier et au second système d'équations de Kolmogorov pour les probabilités de passage.

Soient

$$\lim_{t \downarrow 0} \frac{p_{\{i\}\beta}(t)}{t} = b_{i\beta} \quad (\beta \neq \{i\}), \quad \lim_{t \downarrow 0} \frac{1 - p_{\{i\}\{i\}}(t)}{t} = b_{ii},$$

et supposons que

$$b_{ii} = \sum_{\beta \in M, \, \beta \neq \{i\}} b_{i\beta} < \infty, \quad i = 1, \ldots, n.$$

Les probabilités de passage $p_{\{i\}\beta}(t)$ sont alors solutions du premier système d'équations de Kolmogorov (cf. § 3, (1))

$$\frac{dp_{\{i\}\beta}(t)}{dt} = -b_{ii} p_{\{i\}\beta}(t) + \sum_{\alpha \in M, \, \alpha \neq \{i\}} b_{i\alpha} p_{\alpha\beta}(t).$$

En multipliant cette équation par $s_1^{b_1}, \ldots, s_n^{b_n}$, en sommant sur tous les β et en remarquant que (1) entraîne

$$\sum_{\beta \in M} p_{\alpha\beta}(t) s_1^{b_1} s_2^{b_2} \ldots s_n^{b_n} = \prod_{i=1}^{n} [F_i(t, \sigma)]^{a_i}$$

(ceci exprime l'indépendance mutuelle des évolutions des particules existant à l'instant envisagé), on obtient

$$\frac{\partial F_i(t, \sigma)}{\partial t} = -b_{ii}F_i(t, \sigma) + \sum_{\alpha \in M, \alpha \neq \{i\}} b_{i\alpha} \prod_{i=1}^{n} [F_i(t, \sigma)]^{a_i},$$

ou

$$\frac{\partial F_i(t, \sigma)}{\partial t} = u_i(F_1(t, \sigma), \ldots, F_n(t, \sigma)), \quad i = 1, \ldots, n, \tag{34}$$

où

$$u_i(s_1, \ldots, s_n) = -b_{ii}s_i + \sum_{\alpha \in M, \alpha \neq \{i\}} b_{i\alpha} s_1^{a_1} \ldots s_n^{a_n}, \tag{35}$$

$$i = 1, \ldots, n.$$

Les fonctions $u_i(s_1, \ldots, s_n)$ sont les fonctions génératrices des systèmes de variables $\{-b_{ii}, b_{i\alpha}, \alpha \in M, \alpha \neq \{i\}\}$.

Pour déduire la deuxième équation, on supposera que $|s_i| < 1$ $(i = 1, \ldots, n)$. Alors $|F_i(t, \sigma)| < 1$ $(i = 1, \ldots, n)$ et nous pouvons dériver (33) par rapport à τ. En faisant ensuite $\tau = 0$, on trouve

$$\frac{\partial \Phi(t, \sigma)}{\partial t} = \sum_{k=1}^{n} u_k(\sigma) \frac{\partial \Phi(t, \sigma)}{\partial s_k}. \tag{36}$$

L'équation (36) est un système d'équations identiques pour les fonctions génératrices $F_i(t, \sigma)$:

$$\frac{\partial F_i(t, \sigma)}{\partial t} = \sum_{k=1}^{n} u_k(\sigma) \frac{\partial F_i(t, \sigma)}{\partial s_k}, \quad i = 1, \ldots, n,$$

auquel il faut ajouter les conditions initiales (31). On est donc conduit au

THÉORÈME 7. *Le système de fonctions génératrices* $F_i(t, \sigma)$, $|s_i| < 1$, $i = 1, \ldots, n$, *vérifie le système d'équations différentielles* (34), *l'équation aux dérivées partielles* (36) *et les conditions initiales* (31).

CHAPITRE VIII

PROCESSUS DE DIFFUSION

Dans ce chapitre on se propose d'étudier des processus markoviens continus à valeurs dans un espace euclidien $\mathcal{R}^m$ m-dimensionnel. Ces processus n'ont pas été entièrement décrits à ce jour. On se penchera sur les processus de diffusion qui constituent la plus importante classe. Comme leurs noms l'indiquent, les processus de cette nature peuvent servir de modèle probabiliste à la diffusion. Au § 3, chapitre VI, nous avons examiné le processus du mouvement brownien comme modèle de diffusion en milieu homogène. En reprenant une construction analogue en milieu non homogène, on est conduit à la notion de processus général de diffusion. Précisons ceci sur l'exemple d'un processus à une dimension.

Soit x_t la coordonnée d'une particule assez petite en suspension dans un liquide à l'instant t.

Si l'on néglige l'inertie de la particule, on peut admettre que le déplacement de la particule est la résultante de deux composantes: un déplacement « centré », dû à la vitesse macroscopique du liquide, et les fluctuations provoquées par l'agitation thermique chaotique des molécules du liquide.

Soit $a(t, x)$ vitesse du mouvement macroscopique du liquide au point x à l'instant t. On supposera que la composante fluctuative est une variable aléatoire dont la répartition dépend de la position x de la particule, de l'instant t où est considéré le déplacement, de la durée Δt, de l'intervalle pendant lequel est envisagé le déplacement (on admet que la valeur moyenne de ce déplacement est nulle indépendamment des valeurs prises par t, x_t, Δt). Le déplacement de la particule s'écrit alors approximativement:

$$x_{t+\Delta t} - x_t = a(t, x_t)\,\Delta t + \xi_{t, x_t, \Delta t}\,; \tag{1}$$

$\mathsf{E}\xi_{t, x_t, \Delta t} = 0$. Si $a(t, x)$ est nulle et la répartition de $\xi_{t, x_t, \Delta t}$ ne dépend pas de x comme nous l'avons admis en étudiant le mouvement brownien (§ 3, chapitre VI), alors $\mathsf{E}\xi^2_{t, \Delta t} = \lambda\Delta t$. Le processus est homogène puisqu'il semble naturel d'admettre que de petites variations de t et de x sont pratiquement sans effet sur

les propriétés du milieu ; on supposera donc que $\xi_{t, x_t, \Delta t} = \sigma(t, x) \times \xi_{t, \Delta t}$, où $\sigma(t, x)$ caractérise les propriétés du milieu au point x à l'instant t, et $\xi_{t, \Delta t}$ la valeur de l'accroissement en milieu homogène sachant que $\sigma(t, x) = 1$. Par suite $\xi_{t, \Delta t}$ est égal à l'accroissement du processus du mouvement brownien : $w(t+\Delta t) - w(t)$.

On a donc l'approximation suivante pour l'accroissement $x_{t+\Delta t} - x_t$:

$$x_{t+\Delta t} - x_t \approx a(t, x_t)\,\Delta t + \sigma(t, x_t)\,[w(t+\Delta t) - w(t)]. \qquad (2)$$

Pour que cette formule soit exacte il faut, comme on le fait ordinairement en Analyse, remplacer les accroissements par les différentielles. Ce faisant, on obtient l'équation différentielle

$$dx_t = a(t, x_t)\,dt + \sigma(t, x_t)\,dw(t) \qquad (3)$$

qui peut servir de point de départ à la définition du processus de diffusion.

Si x_t est un processus à plusieurs dimensions, à valeurs dans $\mathcal{R}^m$, la relation (1) reste valable sous réserve que $a(t, x_t)$ soit une fonction à valeurs dans $\mathcal{R}^m$ et $\xi_{t, x_t, \Delta t}$ un vecteur aléatoire dans $\mathcal{R}^m$. On admettra alors que

$$\xi_{t, x_t, \Delta t} = \sum_{k=1}^{m} b_k(t, x_t)\,[w_k(t+\Delta t) - w_k(t)],$$

où $b_k(t, x_t)$ sont des fonctions à valeurs dans $\mathcal{R}^m$ et $w_k(t)$ des processus du mouvement brownien indépendants à une dimension. Cette représentation correspond à un milieu non isotrope : les déplacements dans les divers sens admettent généralement des répartitions différentes. Dans ce cas l'équation en x_t s'écrit

$$dx_t = a(t, x_t)\,dt + \sum_{k=1}^{m} b_k(t, x_t)\,dw_k(t). \qquad (4)$$

A noter que les équations (3) et (4) ne peuvent encore être interprétées de façon stricte. En effet, la quantité $\frac{w(t+\Delta t) - w(t)}{\Delta t}$, si $w(t)$ est un processus du mouvement brownien, admet une répartition normale d'espérance mathématique 0 et de variance $\frac{1}{\Delta t}$. Donc elle ne possède aucune limite probabiliste. Comme $w(t)$ n'a pas de dérivée, la définition usuelle de la différentielle $dw(t)$ n'a pas de sens.

Les équations (3) et (4) seront rigoureusement justifiées après l'introduction au § 1 des notions d'intégrale et de différentielle stochastiques.

§ 1. Intégrale stochastique de Ito

Soient $w(t)$ processus wienérien, $\mathfrak{F}_t$ famille de tribus définies sur l'espace probabilisé fondamental $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ telles que

1) $\mathfrak{F}_{t_1} \subset \mathfrak{F}_{t_2} \subset \mathfrak{S}$ pour $t_1 < t_2$;

2) $w(t)$ soit mesurable par rapport à $\mathfrak{F}_t$, $\forall t$;

3) $w(t+h) - w(h)$, $\forall h$, ne dépende d'aucun événement de $\mathfrak{F}_h$.

Appelons d'autre part $\mathfrak{M}_2[a, b]$ $(0 \leqslant a < b)$ l'ensemble des fonctions $f(t) = f(t, \omega)$ mesurables en (t, ω), définies pour $t \in [a, b]$, $\omega \in \Omega$ et telles que:

a) $f(t)$ soit mesurable par rapport à la tribu $\mathfrak{F}_t$, $\forall t\, [a, b]$;

b) l'intégrale

$$\int_a^b |f(t)|^2\, dt$$

soit presque sûrement finie.

Pour toutes les fonctions de $\mathfrak{M}_2[a, b]$ on définit plus bas l'intégrale $\int_a^b f(t)\, dw(t)$.

Une fonction $f(t)$ est dite *en escalier* si existe une partition de l'intervalle $[a, b]$: $a = t_0 < t_1 < \ldots < t_r = b$, telle que $f(t) = f(t_i)$ pour $t \in [t_i, t_{i+1}[$, $i = 0, \ldots, r-1$.

Commençons par définir l'intégrale $\int_a^b f(t)\, dw(t)$ pour le cas où la fonction $f(t)$ est en escalier. Soit $f(t) = f(t_i)$ pour $t \in [t_i, t_{i+1}[$, où $a = t_0 < t_1 < \ldots < t_r = b$ est une partition de l'intervalle $[a, b]$. Posons

$$\int_a^b f(t)\, dw(t) = \sum_{k=0}^{r-1} f(t_k)\, [w(t_{k+1}) - w(t_k)].$$

Signalons quelques propriétés de cette intégrale.

I. *Si* $\mathsf{E}(|f(t)| \mid \mathfrak{F}_a) < \infty$ *pour* $t \in [a, b]$, *alors*

$$\mathsf{E}\left(\int_a^b f(t)\, dw(t) \mid \mathfrak{F}_a\right) = 0 \tag{1}$$

presque sûrement.

II. *Si* $\mathsf{E}(|f(t)|^2 | \mathfrak{F}_a) < \infty$, *alors*

$$\mathsf{E}\left(\left[\int_a^b f(t)\,dw(t)\right]^2 \Big| \mathfrak{F}_a\right) = \int_a^b \mathsf{E}(f^2(t) | \mathfrak{F}_a)\,dt \pmod{\mathsf{P}} \tag{2}$$

presque sûrement pour $t \in [a, b]$.

En effet,

$$\mathsf{E}\left(\left[\int_a^b f(t)\,dw(t)\right]^2 \Big| \mathfrak{F}_a\right) = \sum_{k=0}^{r-1} \mathsf{E}[f(t_k)^2 (w(t_{k+1}) - w(t_k))^2 | \mathfrak{F}_a] +$$

$$+ 2\sum_{j<k} \mathsf{E}[f(t_j) f(t_k)(w(t_{j+1}) - w(t_j))(w(t_{k+1}) - w(t_k)) | \mathfrak{F}_a] =$$

$$= \sum_{k=0}^{r-1} \mathsf{E}\{f^2(t_k)\,\mathsf{E}[(w(t_{k+1}) - w(t_k))^2 | \mathfrak{F}_{t_k}] | \mathfrak{F}_a\} +$$

$$+ 2\sum_{j<k} \mathsf{E}\{f(t_j) f(t_k)) [w(t_{j+1}) - w(t_j)] \times$$

$$\times \mathsf{E}(w(t_{k+1}) - w(t_k) | \mathfrak{F}_{t_k}) | \mathfrak{F}_a\} =$$

$$= \sum_{k=0}^{r-1} \mathsf{E}(f^2(t_k) | \mathfrak{F}_a)[t_{k+1} - t_k] = \int_a^b \mathsf{E}(f^2(t) | \mathfrak{F}_a)\,dt.$$

III. *Si* $\varphi(t)$ *est une fonction en escalier, alors*

$$\mathsf{P}\left\{\left|\int_a^b \varphi(t)\,dw(t)\right| > c\right\} \leqslant \frac{N}{c^2} + \mathsf{P}\left\{\int_a^b |\varphi(t)|^2\,dt > N\right\},$$

$$\forall N > 0, \ \forall c > 0. \tag{3}$$

En effet, soit $\varphi(t) = \varphi(t_i)$ pour $t_i \leqslant t \leqslant t_{i+1}$, où $a = t_0 < t_1 < \ldots$ $\ldots < t_r = b$. La fonction $\varphi(t)$ étant mesurable par rapport à $\mathfrak{F}_{t_i}$ pour $t \in [t_i, t_{i+1}[$, l'intégrale $\int_a^{t_{i+1}} |\varphi(t)|^2\,dt$ le sera également.

Posons $\varphi_N(t) = \varphi(t)$, $t \leqslant t_i$, pour les i tels que $\int_a^{t_{i+1}} |\varphi(t)|^2\,dt \leqslant$ $\leqslant N$; si pour un k

$$\int_a^{t_k} |\varphi(t)|^2\,dt \leqslant N < \int_a^{t_{k+1}} |\varphi(t)|^2\,dt,$$

alors $\varphi_N(t)=0$ pour $t\in[t_k, b]$. De toute évidence

$$\int_0^t |\varphi_N(t)|^2\,dt\leqslant N$$

et

$$\mathsf{P}\{\sup_t|\varphi_N(t)-\varphi(t)|>0\}=\mathsf{P}\left\{\int_a^b|\varphi(t)|^2\,dt>N\right\}.$$

Donc

$$\mathsf{P}\left\{\left|\int_a^b\varphi(t)\,dw(t)\right|>c\right\}=$$

$$=\mathsf{P}\left\{\left|\int_a^b\varphi_N(t)\,dw(t)+\int_a^b(\varphi(t)-\varphi_N(t))\,dw(t)\right|>c\right\}\leqslant$$

$$\leqslant\mathsf{P}\left\{\left|\int_a^b\varphi_N(t)\,dw(t)\right|>c\right\}+\mathsf{P}\left\{\left|\int_a^b[\varphi(t)-\varphi_N(t)]\,dw(t)\right|>0\right\}\leqslant$$

$$\leqslant\frac{\mathsf{E}\left|\int_a^b\varphi_N(t)\,dw(t)\right|^2}{c^2}+\mathsf{P}\left\{\int_a^b|\varphi(t)|^2\,dt>N\right\}. \quad \blacksquare$$

Soit f_n suite de fonctions en escalier telle que

$$\int_a^b[f(t)-f_n(t)]^2\,dt\to 0$$

stochastiquement. Alors $\int_a^b|f_n(t)-f_m(t)|^2\,dt$ tend aussi stochastiquement vers 0 pour $n\to\infty$, $m\to\infty$. Donc,

$$\lim_{n,\,m\to\infty}\mathsf{P}\left\{\int_a^b|f_n(t)-f_m(t)|^2\,dt>\varepsilon\right\}=0,\quad\forall\varepsilon>0.$$

En utilisant la propriété III, on peut écrire

$$\overline{\lim_{n,\,m\to\infty}}\,\mathsf{P}\left\{\left|\int_a^b f_n(t)\,dw(t)-\int_a^b f_m(t)\,dw(t)\right|>\delta\right\}\leqslant$$

$$\leqslant\frac{\varepsilon}{\delta^2}+\overline{\lim_{n,\,m\to\infty}}\,\mathsf{P}\left\{\int_a^b|f_n(t)-f_m(t)|^2\,dt>\varepsilon\right\}=\frac{\varepsilon}{\delta^2}.$$

$\forall\varepsilon > 0$ et $\forall\delta > 0$, d'où puisque $\varepsilon > 0$ est arbitraire, pour tout $\delta > 0$

$$\lim_{n,\, m\to\infty} \mathsf{P}\left\{\left|\int_a^b f_n(t)\,dw(t) - \int_a^b f_m(t)\,dw(t)\right|^2 > \delta\right\} = 0.$$

La dernière égalité entraîne que la suite de variables aléatoires $\int_a^b f_n(t)\,dw(t)$ converge stochastiquement vers une limite qui ne dépend pas du choix de $f_n(t)$ sous réserve que $\int_a^b |f_n(t) - f(t)|^2\,dt \to 0$ (si existent deux suites $f_n(t)$ et $\bar{f}_n(t)$, en les groupant en une seule on s'assure que leurs limites séquentielles sont presque sûrement égales). Posons

$$\int_a^b f(t)\,dw(t) = \mathsf{P}\text{-}\lim \int_a^b f_n(t)\,dw(t).$$

Cette limite s'appelle *intégrale stochastique de Ito* de la fonction $f(t)$.

L'intégrale stochastique de Ito est définie pour tous les $f \in \mathfrak{M}_2[a, b]$, puisque on peut construire une suite de fonctions en escalier $f_n \in \mathfrak{M}_2[a, b]$ telle que

$$\lim_{n\to\infty} \int_a^b [f_n(t) - f(t)]^2\,dt = 0$$

presque sûrement.

Ainsi l'intégrale définie est une fonctionnelle homogène et additive de la fonction $f(t)$ sur $\mathfrak{M}_2[a, b]$. De plus, pour $a < c < b$

$$\int_a^c f(t)\,dw(t) + \int_c^b f(t)\,dw(t) = \int_a^b f(t)\,dw(t). \tag{4}$$

La démonstration de ces propriétés est évidente pour les fonctions en escalier, elle se transpose au cas général par un trivial passage à la limite.

Dans (3), en passant à la limite sur les fonctions en escalier, on s'assure que la propriété III est valable pour tous les $f(t) \in \mathfrak{M}_2[a, b]$.

Grâce à cette propriété on établit sans peine que

IV. *Si* $f(t) \in \mathfrak{M}_2[a, b]$, $f_n(t) \in \mathfrak{M}_2[a, b]$ *et*

$$\int_a^b |f_n(t) - f(t)|^2\,dt \to 0$$

stochastiquement, alors

$$\mathsf{P}\text{-}\lim \int_a^b f_n(t)\, dw(t) = \int_a^b f(t)\, dw(t).$$

Par passage à la limite on déduit aussitôt la propriété suivante qui généralise les propriétés I et II.

II bis. *Si une fonction f est telle que*

$$\int_a^b \mathsf{E}(|f(t)|^2 | \mathfrak{F}_a)\, dt < \infty$$

presque sûrement, alors

$$\mathsf{E}\left(\int_a^b f(t)\, dw(t) | \mathfrak{F}_a\right) = 0 \quad (\mathrm{mod}\, \mathsf{P}), \tag{5}$$

$$\mathsf{E}\left(\left[\int_a^b f(t)\, dw(t)\right]^2 | \mathfrak{F}_a\right) = \int_a^b \mathsf{E}(|f(t)|^2 | \mathfrak{F}_a)\, dt \quad (\mathrm{mod}\, \mathsf{P}). \tag{6}$$

Traitons maintenant l'intégrale stochastique comme une fonction de la borne supérieure. Soit

$$\psi_t(s) = \begin{cases} 1 & \text{pour } s < t, \\ 0 & \text{pour } s > t. \end{cases}$$

Si $f(s) \in \mathfrak{M}_2[a, b]$, alors $f(s)\psi_t(s) \in \mathfrak{M}_2[a, b]$, $\forall t \in [a, b]$ et

$$\int_a^t f(s)\, dw(s) = \int_a^b f(s)\, \psi_t(s)\, dw(s).$$

Par définition, l'intégrale stochastique est connue aux événements de probabilité nulle près. Donc l'intégrale comme fonction de la borne supérieure est définie à l'équivalence stochastique près (cf. § 1, chapitre IV). Dans toute la suite nous supposerons que les valeurs de l'intégrale comme fonction de la borne supérieure pour tout t sont adaptées de telle sorte que $\zeta(t) = \int_a^t f(s)\, dw(s)$ soit un processus séparable.

Voici les propriétés fondamentales de la fonction $\zeta(t) = \int_a^t f(s)\, dw(s)$.

V. *Si* $\int_a^b \mathsf{E}(|f(s)|^2 | \mathfrak{F}_a)\, ds < \infty$, *alors*

$$\mathsf{P}\left\{\sup_{a\leqslant t\leqslant b}\left|\int_a^t f(s)\,dw(s)\right|>c\,|\,\mathfrak{F}_a\right\}\leqslant\frac{1}{c^2}\int_a^b \mathsf{E}\,(|f(s)|^2\,|\,\mathfrak{F}_a)\,ds \qquad (7)$$

et

$$\mathsf{P}\left\{\sup_{a\leqslant t\leqslant b}\left|\int_a^t f(s)\,dw(s)\right|>c\right\}\leqslant\frac{1}{c^2}\int_a^b \mathsf{E}\,|f(s)|^2\,ds. \qquad (8)$$

Il suffit de prouver (7). Soit une partition de l'intervalle $[a,\ b]:\ a=t_0<t_1<\ldots<t_n=b$. Posons $\zeta_k=\int_a^{t_k} f(s)\,dw(s)$.

Comme pour $k<l$

$$\mathsf{E}\,(\zeta_l-\zeta_k\,|\,\mathfrak{F}_{t_k})=\mathsf{E}\left(\int_{t_k}^{t_l} f(s)\,dw(s)\,|\,\mathfrak{F}_{t_k}\right)=0$$

et ζ_k est mesurable par rapport à $\mathfrak{F}_{t_k}$, la suite $\{\zeta_k\}$ est une martingale et partant $\{\zeta_k^2\}$ est une submartingale. Par suite, en vertu du théorème 5, § 1, chapitre III,

$$\mathsf{P}\{\sup_{0\leqslant k\leqslant n}|\zeta_k|>c\,|\,\mathfrak{F}_a\}\leqslant\frac{1}{c^2}\,\mathsf{E}\,(\zeta_n^2\,|\,\mathfrak{F}_a).$$

Donc

$$\mathsf{P}\left\{\sup_{0\leqslant k\leqslant n}\left|\int_a^{t_k} f(s)\,dw(s)\right|>c\,|\,\mathfrak{F}_a\right\}\leqslant\frac{1}{c^2}\int_a^b \mathsf{E}\,(|\,f(s)|^2\,\mathfrak{F}_a)\,ds,$$

d'où l'on déduit sans peine la démonstration de la propriété V en utilisant la séparabilité du processus $\int_a^t f(s)\,dw(s)$.

VI. *Le processus séparable* $\zeta(t)=\int_a^t f(s)\,dw(s)$ *est continu.*

D é m o n s t r a t i o n. Si $f(t)$ est une fonction en escalier, la continuité de $\zeta(t)$ résulte de celle de $w(t)$ et de la formule qui en l'occurrence définit $\zeta(t)$. Soient $f(t)\in\mathfrak{M}_2\,[a,\ b]$ telle que $\int_a^b \mathsf{E}\,|\,f(s)\,|^2\,ds<\infty$, et $f_n(t)$ suite de fonctions en escalier telle que

$$\lim_{n\to\infty}\int_a^b \mathsf{E}\,|f(s)-f_n(s)|^2\,ds=0.$$

La propriété V nous dit que

$$\mathsf{P}\left\{\sup_t \left| \int_a^t f(s)\, dw(s) - \int_a^t f_n(s)\, dw(s) \right| > \varepsilon \right\} \leqslant$$

$$\leqslant \frac{1}{\varepsilon^2} \int_a^b \mathsf{E}|f(s) - f_n(s)|^2\, ds.$$

En choisissant les suites $\varepsilon_k \to 0$ et n_k telles que

$$\sum_{k=1}^{\infty} \frac{1}{\varepsilon_k^2} \int_a^b \mathsf{E}|f(t) - f_{n_k}(t)|^2\, dt < \infty,$$

on s'assure que

$$\sum_{k=1}^{\infty} \mathsf{P}\left\{ \sup_{a \leqslant t \leqslant b} \left| \int_a^t f(s)\, dw(s) - \int_a^t f_{n_k}(s)\, dw(s) \right| > \varepsilon_k \right\} < \infty,$$

et par suite, d'après le lemme de Borel-Cantelli, à partir d'un certain k, on a presque sûrement

$$\sup_{a \leqslant t \leqslant b} \left| \int_a^t f(s)\, dw(s) - \int_a^t f_{n_k}(s)\, dw(s) \right| \leqslant \varepsilon_k.$$

Donc $\int_a^t f(s)\, dw(s)$ est presque sûrement limite uniforme de fonctions continues: cette limite sera également continue. Soit enfin $f(t)$ fonction arbitraire de $\mathfrak{M}_2[a, b]$. Posons

$$f_N(s) = \begin{cases} f(s) & \text{si } \int_a^s |f(u)|^2\, du \leqslant N, \\ 0 & \text{si } \int_a^s |f(u)|^2\, du > N. \end{cases}$$

On a alors

$$\mathsf{P}\left\{ \sup_{a \leqslant t \leqslant b} \left| \int_a^t f(s)\, dw(s) - \int_a^t f_N(s)\, dw(s) \right| > 0 \right\} \leqslant$$

$$\leqslant \mathsf{P}\left\{ \int_a^b |f(s)|\, ds > N \right\}.$$

Le processus $\int_a^t f(s)\,dw(s)$ est continu dans le cas général, puisque $\int_a^t f_N(s)\,dw(s)$ est continue et la probabilité de droite peut être rendue arbitrairement petite.

On aura besoin de la majoration suivante du moment d'ordre quatre de l'intégrale stochastique:

VII. *Si* $f(t) \in \mathfrak{M}_2[a, b]$ *est telle que* $\int_a^b \mathsf{E}|f(t)|^4\,dt < \infty$, *alors*

$$\mathsf{E}\left(\int_a^b f(t)\,dw(t)\right)^4 \leqslant 36(b-a)\int_a^b \mathsf{E}|f(t)|^4\,dt. \tag{9}$$

D é m o n s t r a t i o n. Supposons tout d'abord que $f(t)$ est une fonction en escalier telle que $f(t) = f(t_i)$ pour $t_i \leqslant t < t_{i+1}$, où $a = t_0 < \ldots < t_r = b$ est une partition de l'intervalle $[a, b]$. Il vient

$$\mathsf{E}\left(\int_a^b f(t)\,dw(t)\right)^4 = \mathsf{E}\left(\sum_{k=0}^{r-1} f(t_k)\,[w(t_{k+1}) - w(t_k)]\right)^4 =$$

$$= \mathsf{E}\sum_{k=0}^{r-1} |f(t_k)|^4\,[w(t_{k+1}) - w(t_k)]^4 +$$

$$+ 6\sum_{k=1}^{r-1} \mathsf{E}\left(\sum_{i=0}^{k-1} f(t_i)\,[w(t_{i+1}) - w(t_i)]\right)^2 |f(t_k)|^2\,[w(t_{k+1}) - w(t_k)]^2 =$$

$$= 3\sum_{k=0}^{r-1} \mathsf{E}|f(t_k)|^4\,(t_{k+1} - t_k)^2 +$$

$$+ 6\sum_{k=1}^{r-1} \mathsf{E}\left(\sum_{i=0}^{k-1} f(t_i)\,[w(t_{i+1}) - w(t_i)]\right)^2 |f(t_k)|^2\,(t_{k+1} - t_k),$$

puisque

$$\mathsf{E}\left([w(t_{k+1}) - w(t_k)]^m / \mathfrak{F}_{t_k}\right) = \begin{cases} 0 \text{ si } m \text{ impair}, \\ (m-1)!!\,(t_{k+1} - t_k)^{m/2} \text{ si } m \text{ pair}. \end{cases}$$

Pour toute fonction en escalier $f(t)$ on peut admettre que les intervalles $[t_k, t_{k+1}]$ sont choisis tels que $\max_k\,[t_{k+1} - t_k]$ soit aussi petit que l'on veut. Donc, en passant à la limite pour

$\max\limits_k [t_{k+1} - t_k] \to 0$ dans l'expression précédente, on obtient

$$\mathsf{E}\left(\int_a^b f(t)\,dw(t)\right)^2 = 6\int_a^b \mathsf{E}\left(\int_a^t f(s)\,dw(s)\right)^2 f^2(t)\,dt. \qquad (10)$$

L'inégalité de Cauchy-Bouniakovski nous dit que

$$\int_a^b \mathsf{E}\left(\int_a^t f(s)\,dw(s)\right)^2 f^2(t)\,dt \leqslant$$

$$\leqslant \left[\int_a^b \mathsf{E}|f(t)|^4\,dt \cdot \int_a^b \mathsf{E}\left(\int_a^t f(s)\,dw(s)\right)^4 dt\right]^{1/2}.$$

De la formule (10) il suit que

$$\mathsf{E}\left(\int_a^t f(u)\,dw(u)\right)^4 = 6\int_a^t \mathsf{E}\left(\int_a^u f(s)\,dw(s)\right)^2 f^2(u)\,du$$

croît avec t. Donc

$$\int_a^b \mathsf{E}\left(\int_a^t f(s)\,dw(s)\right)^4 dt \leqslant (b-a)\,\mathsf{E}\left(\int_a^b f(s)\,dw(s)\right)^4.$$

Par suite

$$\mathsf{E}\left(\int_a^b f(t)\,dw(t)\right)^4 \leqslant 6\left[(b-a)\int_a^b \mathsf{E}|f(s)|^4\,ds\,\mathsf{E}\left(\int_a^b f(s)\,dw(s)\right)^4\right]^{1/2}.$$

D'où la formule (9) pour les fonctions en escalier $f(t)$. On démontre cette formule dans le cas général en construisant pour $f(t)$ une suite de fonctions en escalier $f_n(t)$ telle que

$$\lim_{n\to\infty}\int_a^b \mathsf{E}|f(t) - f_n(t)|^4\,dt = 0.$$

Introduisons la notion de différentielle stochastique. Si pour $\zeta(t)$, processus $\mathfrak{F}_t$-mesurable pour tout t, existent $b(t) \in \mathfrak{M}_2[a, b]$ et $a(t)$ $\mathfrak{F}_t$-mesurable pour tout t et possédant presque sûrement l'intégrale finie $\int_a^b |a(t)|\,dt$, tels que

$$\zeta(t_2) - \zeta(t_1) = \int_{t_1}^{t_2} a(t)\,dt + \int_{t_1}^{t_2} b(t)\,dw(t)$$

pour tous les $a \leqslant t_1 < t_2 \leqslant b$, on dira que $a(t)\,dt + b(t)\,dw(t)$ est la différentielle stochastique de $\zeta(t)$ et l'on notera

$$d\zeta(t) = a(t)\,dt + b(t)\,dw(t).$$

Nous allons établir une propriété très importante de la différentielle stochastique : la *formule de Ito* de différentiation d'une fonction composée.

VIII. *Soient* $\zeta(t)$ *processus possédant la différentielle stochastique* $d\zeta(t) = a(t)\,dt + b(t)\,dw(t)$, $u(t, x)$, *fonction (non aléatoire) définie pour* $t \in [a, b]$, $x \in \mathcal{R}^1$, *continue et possédant des dérivées continues* $u'_t(t, x)$, $u'_x(t, x)$, $u''_{xx}(t, x)$. *Alors le processus* $\eta(t) = u(t, \zeta(t))$ *possède aussi une différentielle stochastique et l'on a la formule suivante appelée formule de Ito :*

$$d\eta(t) = \Big[u'_t(t, \zeta(t)) + u'_x(t, \zeta(t))\,a(t) + \frac{1}{2}u''_{xx}(t, \zeta(t))\,b^2(t)\Big]dt + u'_x(t, \zeta(t))\,b(t)\,dw(t). \qquad (11)$$

Cette formule sera démontrée directement pour le cas multidimensionnel (cf. XIII).

Nous allons étudier des intégrales stochastiques à bornes aléatoires. Soit τ instant markovien par rapport aux tribus $\mathfrak{F}_t$ (i.e. $\{\tau > t\} \in \mathfrak{F}_t$). Supposons que $\tau \in [a, b]$ presque sûrement. Alors pour toute $f(t) \in \mathfrak{M}_2[a, b]$ la fonction $f(t)\,\chi_{\{\tau>t\}} \in \mathfrak{M}_2[a, b]$, puisque $\chi_{\{\tau>t\}}$ $\mathfrak{F}_t$-mesurable.

Posons

$$\int_a^\tau f(t)\,dw(t) = \int_a^b f(t)\,\chi_{\{\tau>t\}}\,dw(t).$$

Si

$$\mathsf{E}\left(\int_a^b f^2(t)\,dt \mid \mathfrak{F}_a\right) < \infty, \qquad (12)$$

alors

$$\mathsf{E}\left(\int_a^\tau f(t)\,dw(t) \mid \mathfrak{F}_a\right) = \mathsf{E}\left(\int_a^b f(t)\,\chi_{\{\tau>t\}}\,dw(t) \mid \mathfrak{F}_a\right) = 0,$$

$$\mathsf{E}\left(\left(\int_a^\tau f(t)\,dw(t)\right)^2 \Big| \mathfrak{F}_a\right) = \mathsf{E}\left(\int_a^b f^2(t)\,\chi_{\{\tau>t\}}\,dt \mid \mathfrak{F}_a\right) = \mathsf{E}\left(\int_a^\tau f^2(t)\,dt \mid \mathfrak{F}_a\right). \qquad (13)$$

Donc, la propriété II bis est valable dans le cas où la borne supérieure de l'intégrale stochastique est instant markovien. On remarque sans peine que la finitude du second membre de (13) équivaut à la réalisation de (12).

Etudions maintenant une intégrale à deux bornes aléatoires. Soit $f(t) \in \mathfrak{M}_2[a, b]$, $\forall b > a$. Si τ_1 et τ_2 sont deux instants markoviens par rapport aux tribus $\mathfrak{F}_t$ tels que $a \leqslant \tau_1 \leqslant \tau_2$ presque sûrement, on posera

$$\int_{\tau_1}^{\tau_2} f(t)\,dw(t) = \int_a^{\tau_2} f(t)\,dw(t) - \int_0^{\tau_1} f(t)\,dw(t).$$

Appelons $\mathfrak{F}_{\tau_1}$ la tribu des événements A de $\bigcup_t \mathfrak{F}_t$ pour lesquels

$$A \cap \{\tau_1 \leqslant t\} \in \mathfrak{F}_t.$$

IX. *Si*

$$\mathsf{E}\left(\int_{\tau_1}^{\tau_2} f(t)\,dw(t)\right) < \infty,$$

alors

$$\mathsf{E}\left(\int_{\tau_1}^{\tau_2} f(t)\,dw(t) \mid \mathfrak{F}_{\tau_1}\right) = 0, \tag{14}$$

$$\mathsf{E}\left(\left[\int_{\tau_1}^{\tau_2} f(t)\,dw(t)\right]^2 \middle| \mathfrak{F}_{\tau_1}\right) = \mathsf{E}\left(\int_{\tau_1}^{\tau_2} f^2(t)\,dt \mid \mathfrak{F}_{\tau_1}\right). \tag{15}$$

D é m o n s t r a t i o n. Soit ξ variable $\mathfrak{F}_{\tau_1}$-mesurable prenant la valeur 0 ou 1. La fonction

$$f_n(t) = \xi\chi_{\{\tau_1 \leqslant t\}} f(t)\,\chi_{\{\tau_2 \wedge n > t\}}$$

appartient à $\mathfrak{M}_2[a, b]$, $\forall b > a$, car $\mathfrak{F}_t$-mesurable (par définition d'une variable $\mathfrak{F}_{\tau_1}$-mesurable on a $\{\omega : \xi = 1\} \cap \{\tau_1 \leqslant t\} \in \mathfrak{F}_t$). Par ailleurs

$$\mathsf{E}\left(\int_a^n f_n^2(t)\,dt\right) = \mathsf{E}\xi \int_{\tau_1 \wedge n}^{\tau_2 \wedge n} f^2(t)\,dt < \infty.$$

Donc

$$\mathsf{E}\xi \int_{\tau_1 \wedge n}^{\tau_2 \wedge n} f(t)\,dw(t) = 0,$$

$$\mathsf{E}\xi \left[\int_{\tau_1 \wedge n}^{\tau_2 \wedge n} f(t)\,dw(t)\right]^2 = \mathsf{E}\xi \int_{\tau_1 \wedge n}^{\tau_2 \wedge n} f(t)\,dw(t).$$

En passant à la limite pour $n \to \infty$ on trouve

$$\mathsf{E}\xi \int_{\tau_1}^{\tau_2} f(t)\, dw(t) = 0,$$

$$\mathsf{E}\xi \left(\int_{\tau_1}^{\tau_2} f(t)\, dw(t) \right)^2 = \mathsf{E}\xi \int_{\tau_1}^{\tau_2} f^2(t)\, dw(t).$$

D'où résultent (14) et (15). ■

Supposons qu'à la famille de tribus $\mathfrak{F}_t$ $(t \geqslant 0)$, définies sur $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ et vérifiant la condition: $\mathfrak{F}_{t_1} \subset \mathfrak{F}_{t_2} \subset \mathfrak{S}$ pour $0 \leqslant t_1 < t_2$, sont associés des processus wienériens à une dimension $w_1(t), \ldots, w_m(t)$ tels que:

a) $w_1(t), \ldots, w_m(t)$ soient indépendants;

b) $w_k(t)$ soit $\mathfrak{F}_t$-mesurable $(t \geqslant 0)$;

c) le processus m-dimensionnel $(w_1(t+h) - w_1(h), \ldots, w_m(t+h) - w_m(h))$, $t \geqslant 0$, ne dépende pas de la tribu $\mathfrak{F}_h$.

Les intégrales

$$\int_a^b f(t)\, dw_k(t)$$

sont définies pour toute fonction $f \in \mathfrak{M}_2[a, b]$. Signalons quelques propriétés des intégrales par rapport aux processus wienériens.

X. *Si $w_1(t)$ et $w_2(t)$ sont indépendants et*

$$\mathsf{E}\left(\int_a^b f^2(t)\, dt \mid \mathfrak{F}_a \right) < \infty, \quad \mathsf{E}\left(\int_a^b g^2(t)\, dt \mid \mathfrak{F}_a \right) < \infty,$$

alors

$$\mathsf{E}\left(\int_a^b f(t)\, dw_1(t) \int_a^b g(t)\, dw_2(t) \mid \mathfrak{F}_a \right) = 0. \tag{16}$$

Supposons d'abord que $f(t)$ et $g(t)$ sont des fonctions en escalier: $f(t) = f(t_k)$, $g(t) = g(t_k)$ pour $t_k \leqslant t < t_{k+1}$, où $a = t_0 < t_1 < \ldots < t_n = b$. On a

$$\mathsf{E}\Big(\sum_{k=0}^{n-1} f(t_k)[w_1(t_{k+1}) - w_1(t_k)] \sum_{j=0}^{n-1} g(t_j)[w_2(t_{j+1}) - w_2(t_j)] \mid \mathfrak{F}_a \Big) =$$

$$= \mathsf{E}\Big(\sum_{k<j} f(t_k) g(t_j)[w_1(t_{k+1}) - w_1(t_k)] \times$$

$$\times \mathsf{E}[w_2(t_{j+1}) - w_2(t_j) \mid \mathfrak{F}_{t_j}] \mid \mathfrak{F}_a \Big) +$$

$$+ \mathsf{E}\Big(\sum_{j<k} f(t_k) g(t_j)[w_2(t_{j+1}) - w_2(t_j)] \times$$

$$\times \mathsf{E}\,[w_1(t_{k+1}) - w_1(t_k) \mid \mathfrak{F}_{t_k}] \mid \mathfrak{F}_a \Big) +$$

$$+ \mathsf{E}\Big(\sum_{k=0}^{n-1} f(t_k)\, g(t_k)\, \mathsf{E}\,([w_1(t_{k+1}) - w_1(t_k)] \times$$

$$\times [w_2(t_{k+1}) - w_2(t_k)] \mid \mathfrak{F}_{t_k})\, \mathfrak{F}_a \Big) = 0.$$

Dans le cas général, (16) se démontre par passage à la limite.

Soit $w(t)$ processus wienérien m-dimensionnel. Définissons l'intégrale

$$\int_a^b (f(t),\ dw(t)) = \sum_{k=1}^{m} \int_a^b f_k(t)\, dw_k(t)$$

pour les fonctions vectorielles $f(t) \in \mathscr{R}^m$ dont les composantes $f_1(t), \ldots, f_m(t) \in \mathfrak{M}_2[a, b]$. Posons pour $x = (x_1, \ldots, x_m)$ de $\mathscr{R}^m$:

$$|x| = \sqrt{x_1^2 + \ldots + x_m^2}.$$

XI. *Si* $f(t)$ *est une fonction telle que* $f_k(t) \in \mathfrak{M}_2[a, b]$ *et* $\mathsf{E}\Big(\int_a^b |f(t)|^2\, dt \mid \mathfrak{F}_a\Big) < \infty$, *alors*

$$\mathsf{E}\Big(\int_a^b (f(t),\ dw(t)) \mid \mathfrak{F}_a\Big) = 0,$$

$$\mathsf{E}\Big(\Big[\int_a^b (f(t),\ dw(t))\Big]^2 \mid \mathfrak{F}_a\Big) = \mathsf{E}\Big(\int_a^b |f(t)|^2\, dt \mid \mathfrak{F}_a\Big).$$

Ces formules découlent des propriétés II bis et X.

Considérons maintenant l'intégrale

$$\int_a^b A(t)\, dw(t), \tag{17}$$

où $A(t)$ est une fonction matricielle: $A(t) = \| a_{ij}(t) \|_{i=1,\ \ldots,\ n}^{j=1,\ \ldots,\ m}$, $a_{ij}(t) \in \mathfrak{M}_2[a, b]$. L'intégrale (17) est une fonction vectorielle à valeurs dans $\mathscr{R}^n$, dont les composantes sont définies par

$$\Big(\int_a^b A(t)\, dw(t)\Big)_i = \sum_{j=1}^{m} \int_a^b a_{ij}(t)\, dw_j(t), \quad i = 1, \ldots, n.$$

En se servant des propriétés II bis et X on établit la propriété suivante.

XII. *Si* $A(t)$ *est une fonction matricielle telle que* $a_{ij}(t) \in \in \mathfrak{M}_2[a, b]$ *et*

$$\mathsf{E}\left(\int_a^b \operatorname{Tr} A(t) A^*(t)\, dt \mid \mathfrak{F}_a\right) = \mathsf{E}\left(\int_a^b \sum_{i,j} a_{ij}^2(t)\, dt \mid \mathfrak{F}_a\right) < \infty,$$

alors

$$\mathsf{E}\left(\left|\int_a^b A(t)\, dw(t)\right|^2 \middle| \mathfrak{F}_a\right) = \mathsf{E}\left(\int_a^b \operatorname{Tr} A(t) A^*(t)\, dt \mid \mathfrak{F}_a\right). \quad (18)$$

Si $B(t) = \|b_{ij}(t)\|_{i=1,\,\dots,\,n}^{j=1,\,\dots,\,m}$, $b_{ij}(t) \in \mathfrak{M}_2[a, b]$ *et* $\mathsf{E}\left(\int_a^b \operatorname{Tr} B(t) \times \right.$ $\left.\times B^*(t)\, dt \mid \mathfrak{F}_a\right) < \infty$, *alors*

$$\mathsf{E}\left[\left(\int_a^b A(t)\, dw(t), \int_a^b B(t)\, dw(t)\right) \middle| \mathfrak{F}_a\right] = \\ = \mathsf{E}\left(\int_a^b \operatorname{Tr} A(t) B^*(t)\, dt \mid \mathfrak{F}_a\right). \quad (19)$$

A gauche, sous le signe de l'espérance mathématique, on reconnaît le produit scalaire de deux vecteurs. (18) est un cas particulier de (19) avec $B(t) = A(t)$. Détaillons le produit scalaire indiqué :

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \int_a^b a_{ij}(t)\, dw_j(t) \sum_{k=1}^{m} \int_a^b b_{ik}(t)\, dw_k(t).$$

Si l'on prend l'espérance mathématique, il ne restera que les termes dont $k = j$. En utilisant la formule II bis, on obtient à droite sous le signe de l'espérance mathématique :

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \int_a^b a_{ij}(t)\, b_{ij}(t)\, dt,$$

ce qui est confondu avec l'expression figurant sous le signe de l'espérance dans le second membre de (19).

Nous allons maintenant généraliser la formule de Ito aux fonctions de plusieurs intégrales stochastiques. On dira qu'un processus $\zeta(t)$ possède une différentielle stochastique sur $[a, b]$ si existent des fonctions $b_k(t) \in \mathfrak{M}_2[a, b]$, $k = 1, \dots, m$, et une fonction

$a(t)$ $\mathfrak{F}_t$-mesurable pour laquelle $\int_a^b |a(t)|\,dt < \infty$, telles que

$$\zeta(t_2)-\zeta(t_1)=\int_{t_1}^{t_2} a(s)\,ds+\sum_{k=1}^{m}\int_{t_1}^{t_2} b_k(s)\,dw_k(s)$$

pour $a \leqslant t_1 < t_2 \leqslant b$. On écrira

$$d\zeta(t)=a(t)\,dt+\sum_{k=1}^{m} b_k(t)\,dw_k(t).$$

XIII. *Si des processus $\zeta_1(t), \ldots, \zeta_n(t)$ possèdent des différentielles stochastiques sur $[a, b]$:*

$$d\zeta_k(t)=a_k(t)\,dt+\sum_{j=1}^{m} b_{kj}(t)\,dw_j(t), \quad k=1, \ldots, n, \tag{20}$$

et si $u(t, x_1, \ldots, x_n)$ est une fonction continue possédant des dérivées continues $\frac{\partial u}{\partial t}$, $\frac{\partial u}{\partial x_k}$, $k=1, \ldots, n$; $\frac{\partial^2 u}{\partial x_i\,\partial x_j}$, $i, j=1, \ldots, n$, alors la fonction $\eta(t)=u(t, \zeta_1(t), \ldots, \zeta_n(t))$ possède aussi une différentielle stochastique:

$$d\eta(t)=\Big[\frac{\partial u}{\partial t}(t, \zeta_1(t), \ldots, \zeta_n(t))+\sum_{k=1}^{n}\frac{\partial u}{\partial x_k}(t, \zeta_1(t), \ldots, \zeta_n(t))\,a_k(t)+$$

$$+\frac{1}{2}\sum_{i,k=1}^{n}\frac{\partial^2 u}{\partial x_i\,\partial x_k}(t, \zeta_1(t), \ldots, \zeta_m(t))\sum_{j=1}^{m} b_{ij}(t)\,b_{kj}(t)\Big]\,dt+$$

$$+\sum_{j=1}^{m}\Big(\sum_{k=1}^{n}\frac{\partial u}{\partial x_k}(t, \zeta_1(t), \ldots, \zeta_n(t))\,b_{kj}(t)\Big)\,dw_j(t). \tag{21}$$

Démonstration. Supposons que les fonctions a_k et b_{kj} sont constantes sur l'intervalle $[s_1, s_2]$ et que la fonction $u(t, x_1, \ldots, x_n)$ est non nulle seulement pour $|x| \leqslant N$. Soit $s_1 = t_0 < \ldots < t_l = s_2$. En utilisant la formule

$$u(t_{k+1}, x_1, \ldots, x_n)-u(t_k, y_1, \ldots, y_n)=$$

$$=\frac{\partial u}{\partial t}(t_k, y_1, \ldots, y_n)[t_{k+1}-t_k]+\sum_{j=1}^{n}\frac{\partial u}{\partial x_j}(t_k, y_1, \ldots, y_n)[x_j-y_j]+$$

$$+\frac{1}{2}\sum_{i,\,j=1}^{n}\frac{\partial^2 u}{\partial x_i\,\partial x_j}(t_k,\,y_1,\,\ldots,\,y_n)(x_i-y_i)(x_j-y_j)+$$

$$+\varepsilon\left[(t_{k+1}-t_k)+\sum_{j=1}^{n}(x_j-y_j)^2\right],$$

où $\varepsilon=\varepsilon(t_k,\,t_{k+1},\,x_1,\,\ldots,\,x_n,\,y_1,\,\ldots,\,y_n)$ tend uniformément vers zéro avec $|\,t_{k+1}-t_k\,|+\sum_{j=1}^{n}[x_j-y_j]^2$, on trouve

$$\eta(t_{k+1})-\eta(t_k)=\beta(t_k)\,\Delta t_k+$$

$$+\sum_{j=1}^{n}\frac{\partial u}{\partial x_j}(t_k,\,\zeta_1(t_k),\,\ldots,\,\zeta_n(t_k))\sum_{i=1}^{m}b_{ji}\Delta w_i(t_k)+$$

$$+\frac{1}{2}\sum_{i,\,j=1}^{n}\frac{\partial^2 u}{\partial x_i\,\partial x_j}(t_k,\,\zeta_1(t_k),\,\ldots,\,\zeta_n(t_k))\times$$

$$\times\left[\Delta\zeta_i(t_h)^2\,\Delta\zeta_j(t_k)-\sum_{r=1}^{m}b_{ir}b_{jr}\Delta t_k\right]+\varepsilon\left[\Delta t_k+\sum_{j=1}^{n}\Delta\zeta_j(t_k)^2\right],\qquad(22)$$

où

$$\beta(t)=\frac{\partial u}{\partial t}(t,\,\zeta_1(t),\,\ldots,\,\zeta_n(t))+\sum_{j=1}^{n}\frac{\partial u}{\partial t}(t,\,\zeta_1(t),\,\ldots,\,\zeta_n(t))\,a_j+$$

$$+\frac{1}{2}\sum_{i,\,j=1}^{n}\frac{\partial^2 u}{\partial x_i\,\partial x_j}(t,\,\zeta_1(t),\,\ldots,\,\zeta_n(t))\sum_{r=1}^{m}b_{ir}b_{jr},$$

$$\Delta t_k=t_{k+1}-t_k,\quad \Delta w_i(t_k)=w_i(t_{k+1})-w_i(t_k),$$

$$\Delta\zeta_j(t_k)=\zeta_j(t_{k+1})-\zeta_j(t_k).$$

Supposons que $\max\limits_{k}\Delta t_k\to 0$. Alors $\max\limits_{k,\,i}|\,\Delta w_i(t_k)\,|\to 0$, $\max\limits_{k,\,i}|\,\Delta\zeta_i(t_k)\,|\to 0$. Donc

$$\lim_{\max\Delta t_k\to 0}\sum\varepsilon\left[\Delta t_k+\sum_{j=1}^{n}\Delta\zeta_j(t_k)^2\right]\to 0$$

stochastiquement, puisque pour un c

$$\sum_{k=0}^{l-1}\Delta\zeta_j(t_k)^2\leqslant c\sum_{k=0}^{l-1}\left[\Delta t_k^2+\sum_{i=1}^{m}\Delta w_i(t_k)^2\right]$$

et le second membre est stochastiquement borné, puisque

$$\mathsf{E}\sum_{k=0}^{l-1} \Delta w_i(t_k)^2 \leqslant s_2 - s_1.$$

Si donc $\max_k \Delta t_k \to 0$, alors

$$\sum_{k=0}^{l-1} \beta(t_k)\,\Delta t_k \to \int_{s_1}^{s_2} \beta(t)\,dt,$$

$$\sum_{k=0}^{l-1} \sum_{j=1}^{n} \frac{\partial u}{\partial x_j}(t_k, \zeta_1(t_k), \ldots, \zeta_n(t_k)) \sum_{i=1}^{m} b_{ji}\Delta w_j(t_k) \to$$

$$\to \sum_{i=1}^{m} \int_{s_1}^{s_2} \sum_{j=1}^{n} \frac{\partial u}{\partial x_j}(t, \zeta_1(t), \ldots, \zeta_n(t))\, b_{ij}\, dw_j(t).$$

Ensuite

$$\mathsf{E}\Big[\Big(\sum_{k=0}^{l-1} \sum_{i,\,j=1}^{n} \frac{\partial^2 u}{\partial x_i\,\partial x_j}(t_k, \zeta_1(t_k), \ldots, \zeta_n(t_k)) \times$$

$$\times \Big\{\Big[a_i\Delta t_k + \sum_{r=1}^{m} b_{ir}\Delta w_r(t_k)\Big] \times$$

$$\times \Big[a_j\Delta t_k + \sum_{r=1}^{m} b_{jr}\Delta w_r(t_k)\Big] - \sum_{r=1}^{m} b_{ir}b_{jr}\Delta t_k\Big\}\Big)^2 \Big| \mathfrak{F}_{s_1}\Big] \leqslant$$

$$\leqslant 2\mathsf{E}\Big[\Big(\sum_{k=0}^{l-1} \sum_{i,\,j=1}^{n} \frac{\partial^2 u}{\partial x_i\,\partial x_j}(t_k, \zeta_1(t_k), \ldots, \zeta_n(t_k))\, a_i^2\,\Delta t_k^2\Big)^2 \Big| \mathfrak{F}_{s_1}\Big] +$$

$$+ 2\mathsf{E}\Big[\Big(\sum_{k=0}^{l-1} \sum_{i,\,j=1}^{n} \frac{\partial^2 u}{\partial x_i\,\partial x_j}(t_k, \zeta_1(t_k), \ldots, \zeta_n(t_k)) \times$$

$$\times \Big[\sum_{r,\,q=1}^{m} b_{ir}b_{jq}\Delta w_r(t_k)\,\Delta w_q(t_k) - \sum_{r=1}^{m} b_{ir}b_{jr}\Delta t_k\Big]\Big)^2 \Big| \mathfrak{F}_{s_1}\Big].$$

Soit $\max_{t,\,x_1,\,\ldots,\,x_n} \max_{i,\,j} \left| \frac{\partial^2 u}{\partial x_i\,\partial x_j}(t, x_1, \ldots, x_n) \right| \leqslant L$. La première somme de gauche est alors majorée par

$$2\Big(nL\sum_{i=1}^{n} a_i^2\Big)^2 \Big(\sum_{k=0}^{l-1} \Delta t_k^2\Big)^2 \leqslant 2\Big(nL\sum_{i=1}^{n} a_i^2\Big)^2 (s_2 - s_1)^2 (\max_k \Delta t_k)^2 \to 0$$

pour $\max_k \Delta t_k \to 0$. On remarquera que les termes de la seconde somme ne sont pas corrélés pour divers k. Cette somme est donc majorée par

$$2\mathsf{E}\Big(\sum_{k=0}^{l-1}\Big(\sum_{i,\,j=1}^{n}\frac{\partial^2 u}{\partial x_i\,\partial x_j}(t_k,\zeta_1(t_k),\ldots,\zeta_n(t_k))\times$$

$$\times\Big[\sum_{r,\,q=1}^{m} b_{ir}b_{jq}\Delta w_r(t_k)\,\Delta w_q(t_k)-\sum_{r=1}^{m} b_{ir}b_{jr}\Delta t_k\Big]\Big)^2\Big|\mathfrak{F}_{s_1}\Big)\leqslant$$

$$\leqslant 2\sum_{k=0}^{l-1} n^2L^2\sum_{i,\,j=1}^{n}\mathsf{E}\Big(\Big[\sum_{r,\,q=1}^{m} b_{ir}b_{jq}\Delta w_r(t_k)\,\Delta w_q(t_k)-$$

$$-\sum_{r=1}^{m} b_{ir}b_{jr}\Delta t_k\Big]^2\Big|\mathfrak{F}_{s_1}\Big)\leqslant$$

$$\leqslant 2n^2L^2\sum_{k=0}^{l-1}\sum_{i,\,j=1}^{n}\Big(2\sum_{r=1}^{m} b_{ir}^2 b_{jr}^2+\sum_{\substack{r\neq q\\ r,\,q=1}}^{m} b_{ir}^2 b_{jq}^2\Big)\Delta t_k^2\to 0$$

pour $\max_k \Delta t_k \to 0$. (A noter que a_k et b_{ij} sont $\mathfrak{F}_{s_1}$-mesurables, car confondues avec $a_k(s_1)$ et $b_{ij}(s_1)$ respectivement.)

On a donc établi la formule (21) pour les fonctions $u(t, x_1, \ldots, x_n)$ à support borné en x et pour les constantes a_j et b_{ij}. Par suite, cette formule est valable pour les fonctions en escalier. Ecrivons-la sous la forme intégrale

$$u(s_2,\zeta_1(s_2),\ldots,\zeta_n(s_2))-u(s_1,\zeta_1(s_1),\ldots,\zeta_n(s_1))=$$

$$=\int_{s_1}^{s_2}\Big[\frac{\partial u}{\partial t}(t,\zeta_1(t),\ldots,\zeta_n(t))+\sum_{k=1}^{n}\frac{\partial u}{\partial x_k}(t,\zeta_1(t),\ldots,\zeta_n(t))\,a_k(t)+$$

$$+\frac{1}{2}\sum_{k,\,j=1}^{n}\frac{\partial^2 u}{\partial x_k\,\partial x_j}(t,\zeta_1(t),\ldots,\zeta_n(t))\sum_{i=1}^{m} b_{ki}(t)\,b_{ji}(t)\Big]dt+$$

$$+\sum_{i=1}^{m}\int_{s_1}^{s_2}\sum_{k}\frac{\partial u}{\partial x_k}(t,\zeta_1(t),\ldots,\zeta_n(t))\,b_{ki}(t)\,dw_i(t). \qquad (23)$$

Dans cette formule on peut passer à la limite sur les fonctions en escalier par rapport à $a_k(t)$ et $b_{kj}(t)$ et ensuite sur les fonctions à support borné $u(t, x_1, \ldots, x_n)$ vérifiant XIII, en se servant de la propriété IV.

§ 2. Existence et unicité des solutions des équations différentielles stochastiques

Soit l'équation différentielle stochastique

$$d\xi(t) = a(t, \xi(t))\,dt + \sigma(t, \xi(t))\,dw(t), \tag{1}$$

dont on supposera naturellement que la solution est un processus de diffusion de coefficient de diffusion $\sigma^2(t, x)$ et de coefficient de transfert $a(t, x)$. On admet que $a(t, x)$ et $\sigma(t, x)$ sont des fonctions boréliennes définies pour $x \in \mathscr{R}^1$ et $t \in [t_0, T]$.

L'équation (1) équivaut à la suivante:

$$\xi(t) = \xi(t_0) + \int_{t_0}^{t} a(s, \xi(s))\,ds + \int_{t_0}^{t} \sigma(s, \xi(s))\,dw(s), \tag{2}$$

et admet une solution sous réserve que $\xi(t_0)$ soit donnée. Pour que les intégrales de (2), et par suite les différentielles de (1), aient un sens, il faut introduire des tribus d'événements $\mathfrak{F}_t$.

Dans toute la suite on supposera que $\xi(t_0)$ est indépendante du processus $w(t) - w(t_0)$ et que $\mathfrak{F}_t$ est la plus petite tribu par rapport à laquelle sont mesurables $\xi(t_0)$ et $w(s) - w(t_0)$ pour $t_0 < s \leqslant t$. On dira qu'un processus $\xi(t)$ est solution de l'équation (2) si $\xi(t)$ est $\mathfrak{F}_t$-mesurable, existent les intégrales de (2) et enfin (2) est réalisée presque sûrement pour tout $t \in [t_0, T]$.

A noter que de la propriété III (§ 1) il résulte que deux processus $f_1(s)$ et $f_2(s)$ stochastiquement équivalents possèdent des intégrales stochastiques

$$\int_{t_0}^{t} f_1(s)\,dw(s), \quad \int_{t_0}^{t} f_2(s)\,dw(s)$$

presque sûrement égales, puisque $f_1(s) = f_2(s)$ presque sûrement pour tout s et par suite

$$\mathsf{P}\left\{\int_{t_0}^{T} |f_1(s) - f_2(s)|^2\,ds > 0\right\} = 0.$$

D'où il suit que tout processus stochastiquement équivalent à une solution de (2) sera lui-même solution de (2). Le second membre de (2) étant stochastiquement équivalent au membre de gauche et continu presque sûrement, toute solution de (2) est stochastiquement équivalente à une solution continue de (2). Dans la suite on n'envisagera que les solutions continues de (2).

THÉORÈME 1. *Soient* $a(t, x)$ *et* $\sigma(t, x)$ $(t \in [t_0, T],\ x \in \mathscr{R}^1)$ *fonctions boréliennes telles que pour un* K

a) $$|a(t, x) - a(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leqslant K|x-y|, \ \forall x, \forall y \in \mathscr{R}^1;$$

b) $$|a(t, x)|^2 + |\sigma(t, x)|^2 \leqslant K^2(1+x^2), \ \forall x \in \mathscr{R}^1.$$

L'équation (2) *admet une solution. Si* $\xi_1(t)$ *et* $\xi_2(t)$ *sont deux solutions continues (pour* $\xi(t_0)$ *fixe), alors*

$$\mathsf{P}\{\sup_{t_0 \leqslant t \leqslant T} |\xi_1(t) - \xi_2(t)| > 0\} = 0.$$

D é m o n s t r a t i o n. Prouvons d'abord l'unicité de la solution continue. Soient $\xi_1(t)$ et $\xi_2(t)$ deux solutions continues de (2). Soit $\chi_N(t)$ variable aléatoire telle que:

$$\chi_N(t) = \begin{cases} 1 & \text{si } |\xi_1(s)| \leqslant N,\ |\xi_2(s)| \leqslant N \text{ pour } s \in [t_0, t]; \\ 0 & \text{dans le cas contraire.} \end{cases}$$

Comme $\chi_N(t)\,\chi_N(s) = \chi_N(t)$ pour $s < t$, on a

$$\chi_N(t)[\xi_1(t) - \xi_2(t)] = \chi_N(t)\Big[\int_{t_0}^{t} \chi_N(s)[a(s, \xi_1(s)) - a(s, \xi_2(s))]\,ds + \int_{t_0}^{t} \chi_N(s)[\sigma(s, \xi_1(s)) - \sigma(s, \xi_2(s))]\,dw(s)\Big].$$

Les espérances mathématiques des carrés des intégrales du second membre existent puisque

$$\chi_N(s)[|a(s, \xi_1(s)) - a(s, \xi_2(s))| + |\sigma(s, \xi_1(s)) - \sigma(s, \xi_2(s))|] \leqslant \leqslant K\chi_N(s)|\xi_1(s) - \xi_2(s)| \leqslant 2KN.$$

L'inégalité $(a+b)^2 \leqslant 2a^2 + 2b^2$, l'inégalité de Cauchy et la priéte II bis du § 2 nous conduisent à l'inégalité

$$\mathsf{E}\chi_N(t)[\xi_1(t) - \xi_2(t)]^2 \leqslant \leqslant 2\mathsf{E}\chi_N(t)\Big(\int_{t_0}^{t} \chi_N(s)[a(s, \xi_1(s)) - a(s, \xi_2(s))]\,ds\Big)^2 + 2\mathsf{E}\chi_N(t)\Big(\int_{t_0}^{t} \chi_N(s)[\sigma(s, \xi_1(s)) - \sigma(s, \xi_2(s))]\,dw(s)\Big)^2 \leqslant$$

$$\leqslant 2(T-t_0)\int_{t_0}^{t} \mathsf{E}\chi_N(s)\,[a(s,\ \xi_1(s))-a(s,\ \xi_2(s))]^2\,ds+$$

$$+2\int_{t_0}^{t} \mathsf{E}\chi_N(s)\,[\sigma(s,\ \xi_1(s))-\sigma(s,\ \xi_2(s))]^2\,ds.$$

La condition a) implique l'existence d'une constante L telle que

$$\mathsf{E}\chi_N(t)\,[\xi_1(t)-\xi_2(t)]^2\leqslant L\int_{t_0}^{t} \mathsf{E}\chi_N(s)\,[\xi_1(s)-\xi_2(s)]^2\,ds. \qquad (3)$$

Nous allons nous servir d'une proposition auxiliaire qui nous sera utile dans de nombreuses estimations.

LEMME 1. *Si $\alpha(t)$ est une fonction intégrable non négative définie pour $t\in[t_0,\ T]$ et vérifiant l'inégalité*

$$\alpha(t)\leqslant H\int_{t_0}^{t}\alpha(s)\,ds+\beta(t), \qquad (4)$$

où H est une constante non négative, et $\beta(t)$ une fonction intégrable, alors

$$\alpha(t)\leqslant\beta(t)+H\int_{t_0}^{t}e^{H(t-s)}\beta(s)\,ds.$$

Démonstration. De (4) il résulte

$$\alpha(t)\leqslant\beta(t)+H\int_{t_0}^{t}\Big[\beta(t)+H\int_{t_0}^{s}\alpha(s_1)\,ds_1\Big]\,ds\leqslant$$

$$\leqslant\beta(t)+H\int_{t_0}^{t}\Big[\beta(s_1)+H\int_{t_0}^{s_1}\Big[\beta(s_2)+H\int_{t_0}^{s_2}\alpha(s_3)\,ds_3\Big]\,ds_2\Big]\,ds_1\leqslant$$

$$\leqslant\beta(t)+H\int_{t_0}^{t}\beta(s_1)\,ds+H^2\int_{t_0}^{t}\int_{t_0}^{s_1}\beta(s_2)\,ds_2\,ds_1+\ \dots$$

$$\dots+H^n\int_{t_0}^{t}\int_{t_0}^{s_1}\dots\int_{t_0}^{s_{n-1}}\beta(s_n)\,ds_1\dots ds_n+$$

$$+H^{n+1}\int_{t_0}^{t}\dots\int_{t_0}^{s_n}\alpha(s_{n+1})\,ds_1\dots ds_{n+1}.$$

Comme

$$\int_{t_0}^{t}\int_{t_0}^{s_1}\cdots\int_{t_0}^{s_n}\alpha(s_{n+1})\,ds_1\ldots ds_{n+1}=\int_{t_0}^{t}\frac{(t-s_{n+1})^n}{n!}\alpha(s_{n+1})\,ds_{n+1},$$

il vient

$$\lim_{n\to\infty}H^{n+1}\int_{t_0}^{t}\cdots\int_{t_0}^{s_n}\alpha(s_{n+1})\,ds_1\ldots ds_{n+1}=0;$$

donc

$$\alpha(t)\leqslant\beta(t)+H\int_{t_0}^{t}\beta(s)\,ds+H^2\int_{t_0}^{t}\int_{t_0}^{s_1}\beta(s_2)\,ds_2\,ds_1+\ldots=$$

$$=\beta(t)+H\int_{t_0}^{t}\sum_{k=0}^{\infty}H^k\frac{(t-s)^k}{k!}\beta(s)\,ds. \blacksquare$$

Si l'on pose

$$\alpha(t)=\mathsf{E}\chi_N(t)\,[\xi_1(t)-\xi_2(t)]^2;\ \beta(t)=0,$$

on déduit de (3)

$$\mathsf{E}\chi_N(t)\,[\xi_1(t)-\xi_2(t)]^2=0,$$

i. e.

$$\mathsf{P}\{\xi_1(t)\neq\xi_2(t)\}\leqslant\mathsf{P}\{\sup_s|\xi_1(s)|>N\}+\mathsf{P}\{\sup_s|\xi_2(s)|>N\}.$$

Les probabilités de droite tendent vers zéro, car les processus $\xi_1(t)$ et $\xi_2(t)$ sont presque sûrement continus (et par suite bornés). Donc $\xi_1(t)$ et $\xi_2(t)$ sont stochastiquement équivalents. Ces processus étant presque sûrement continus, on a $\mathsf{P}\{\sup_t|\xi_1(t)-\xi_2(t)|>0\}=0$. D'où l'unicité de la solution de (2).

Démontrons maintenant l'existence de la solution de (2). Supposons d'abord que $\mathsf{E}|\xi(t_0)|^2<\infty$. Soit $\mathscr{B}$ espace de Banach de fonctions aléatoires $\mathfrak{F}_t$-mesurables pour tout t, vérifiant la relation $\sup_{t_0\leqslant t\leqslant T}\mathsf{E}|\zeta(t)|^2<\infty$ et de norme

$$\|\zeta\|=(\sup_{t\leqslant_0 t\leqslant T}\mathsf{E}|\zeta(t)|^2)^{1/2}.$$

Soit défini sur $\mathscr{B}$ l'opérateur S:

$$S\zeta(t)=\xi(t_0)+\int_{t_0}^{t}a(s,\zeta(s))\,ds+\int_{t_0}^{t}\sigma(s,\zeta(s))\,dw(s).$$

L'existence des deux intégrales est une conséquence de la relation

$$|a(s,\zeta(s))|^2+|\sigma(s,\zeta(s))|^2\leqslant K^2(1+|\zeta(s)|^2).$$

De toute évidence, $S\zeta(t)$ est $\mathfrak{F}_t$-mesurable.

L'inégalité $(a+b+c)^3 \leqslant 3(a^2+b^2+c^2)$ et la condition b) du théorème donnent

$$\mathsf{E}\,|S\zeta(t)|^2 \leqslant 3\mathsf{E}\,|\xi(t_0)|^2 + \\ +3(T-t_0)\,\mathsf{E}\int_{t_0}^{t} K^2(1+|\zeta(s)|^2)\,ds + 3\int_{t_0}^{t} \mathsf{E}K^2(1+|\zeta(s)|^2)\,ds \leqslant \\ \leqslant 3\mathsf{E}\,|\xi(t_0)|^2 + [3(T-t_0)^2K^2 + 3(T-t_0) + K^2](1+\|\zeta\|^2).$$

Donc l'opérateur S applique $\mathscr{B}$ dans $\mathscr{B}$. On a ensuite

$$\mathsf{E}\,|S\zeta_1(t) - S\zeta_2(t)|^2 \leqslant \\ \leqslant 2(T-t_0)\int_{t_0}^{t} \mathsf{E}\,[a(s,\zeta_1(s)) - a(s,\zeta_2(s))]^2\,ds + \\ +2\mathsf{E}\left(\int_{t_0}^{t} [\sigma(s,\zeta_1(s)) - \sigma(s,\zeta_2(s))]\,dw(s)\right)^2 \leqslant \\ \leqslant L\int_{t_0}^{t} \mathsf{E}\,|\zeta_1(s) - \zeta_2(s)|^2\,ds \leqslant L(t-t_0)\,\|\zeta_1 - \zeta_2\|^2, \\ L = 2K^2(T-t_0+1).$$

Cette relation montre que l'opérateur S est continu sur $\mathscr{B}$. Puis

$$\mathsf{E}\,|S^n\zeta_1(t) - S^n\zeta_2(t)|^2 \leqslant \\ \leqslant L\int_{t_0}^{t} \mathsf{E}\,|S^{n-1}\zeta_1(u) - S^{n-1}\zeta_2(u)|^2\,du \leqslant \\ \leqslant L^n \int\limits_{t_0<t_1<\ldots<t_n<t}\!\!\!\ldots\!\int \mathsf{E}\,|\zeta_1 - \zeta_2|^2\,dt_1\ldots dt_n = \frac{L^n(t-t_0)^n}{n!}\,\|\zeta_1 - \zeta_2\|^2.$$

Donc pour tout $\zeta(t) \in \mathscr{B}$ on a

$$\|S^{n+1}\zeta - S^n\zeta\|^2 \leqslant \frac{L^n(T-t_0)^n}{n!}\,\|S\zeta - \zeta\|^2.$$

La convergence de la série

$$\sum_{n=1}^{\infty} \|S^{n+1}\zeta - S^n\zeta\|$$

entraîne l'existence de la limite du processus $S^n\zeta(t)$ pour $n \to \infty$. Si l'on note cette limite $\xi(t)$, alors de la continuité de S il suit que $S[S^n\zeta(t)] \to S\xi(t)$. Or $S[S^n\zeta(t)] = S^{n+1}\zeta(t) \to \xi(t)$. Donc $\|S\xi - \xi\| = 0$. De la définition de la norme il résulte que $\xi(t) =$

$= S\xi(t)$ presque sûrement pour tout $t \in [t_0, T]$, i.e. $\xi(t)$ est solution de l'équation (2).

Voyons maintenant la démonstration de l'existence de la solution de (2) dans le cas général.

Soit

$$\xi^N(t_0) = \begin{cases} \xi(t_0) & \text{pour } |\xi(t_0)| \leqslant N, \\ 0 & \text{pour } |\xi(t_0)| > N. \end{cases}$$

Notons $\xi^N(t)$ la solution de l'équation

$$\xi^N(t) = \xi^N(t_0) + \int_{t_0}^{t} a(s, \xi^N(s))\, ds + \int_{t_0}^{t} \sigma(s, \xi^N(s))\, dw(s). \tag{5}$$

Comme $|\xi^N(t_0)| \leqslant N$ et $\mathsf{E}|\xi^n(t_0)|^2 < \infty$, l'équation (5) admet une solution et de ce qui a été démontré il suit que $\sup_t \mathsf{E}|\xi^N(t)|^2 < \infty$.

Montrons que $\xi^N(t)$ pour $N \to \infty$ tend stochastiquement vers un processus $\xi(t)$, solution de l'équation (2). Soient $N' > N$ et η une variable aléatoire telle que

$$\eta = \begin{cases} 1 & \text{pour } |\xi(t_0)| \leqslant N, \\ 0 & \text{pour } |\xi(t_0)| > N. \end{cases}$$

Alors $|\xi^N(t_0) - \xi^{N'}(t_0)|\,\eta = 0$. La variable η est $\mathfrak{F}_{t_0}$-mesurable. La relation

$$|\xi^N(t) - \xi^{N'}(t)|^2\,\eta \leqslant 2\left[\int_{t_0}^{t} [a(s, \xi^N(s)) - a(s, \xi^{N'}(s))]\,\eta\, ds\right]^2 + \\ + 2\left(\int_{t_0}^{t} [\sigma(s, \xi^N(s)) - \sigma(s, \xi^{N'}(s))]\,\eta\, dw(s)\right)^2,$$

la condition a) du théorème et les majorations des intégrales nous disent qu'il existe une constante L telle que

$$\mathsf{E}|\xi^N(t) - \xi^{N'}(t)|^2\,\eta \leqslant L \int_{t_0}^{t} \mathsf{E}|\xi^N(s) - \xi^{N'}(s)|^2\,\eta\, ds.$$

Donc

$$\mathsf{E}|\xi^N(t) - \xi^{N'}(t)|^2\,\eta = 0,$$

de sorte que

$$\mathsf{P}\{|\xi^N(t) - \xi^{N'}(t)| > 0\} < \mathsf{P}\{|\xi(t_0)| > N\}.$$

D'où il résulte que $\xi^N(t)$ converge stochastiquement vers une limite $\xi(t)$ pour $N \to \infty$ et de plus

$$\int_{t_0}^{T} (\xi^N(t) - \xi(t))^2 \, dt$$

converge stochastiquement vers 0 pour $N \to \infty$. De la condition a) et de la propriété IV, § 1, il suit que

$$\int_{t_0}^{t} a(s, \xi^N(s)) \, ds \to \int_{t_0}^{t} a(s, \xi(s)) \, ds, \quad \int_{t_0}^{t} \sigma(s, \xi^N(s)) \, dw(s) \to \int_{t_0}^{t} \sigma(s, \xi(s)) \, dw(s)$$

stochastiquement pour $N \to \infty$. Donc $\xi(s)$ est solution de (2). ■

Montrons que dans les hypothèses du théorème 1 la solution de l'équation (2) sera processus markovien. Pour cela nous allons prouver le théorème suivant.

THÉORÈME 2. *Supposons que* $a(t, x)$ *et* $\sigma(t, x)$ *vérifient les hypothèses du théorème* 1 *et que* $\xi_{t,x}(s)$, *processus défini pour* $s \in [t, T]$, $t > t_0$, *est solution de l'équation*

$$\xi_{t,x}(s) = x + \int_{t}^{s} a(u, \xi_{t,x}(u)) \, du + \int_{t}^{s} \sigma(u, \xi_{t,x}(u)) \, dw(u). \tag{6}$$

Alors le processus $\xi(t)$, *solution de l'équation* (2), *sera un processus markovien dont les probabilités de passage sont définies par* $\mathsf{P}(t, x, s, A) = \mathsf{P}\{\xi_{t,x}(s) \in A\}$.

D é m o n s t r a t i o n. $\xi_{t,x}(s)$ ne dépend pas de $\xi(t)$ et des événements de $\mathfrak{F}_t$ puisque $\xi(t)$ est $\mathfrak{F}_t$-mesurable et $\xi_{t,x}(s)$ entièrement défini par le processus $w(s) - w(t)$ pour $s \in [t, T]$ (qui ne dépend pas de $\mathfrak{F}_t$). Le théorème 1 nous dit que $\xi(s)$ pour $s \in [t, T]$ est l'unique solution de l'équation

$$\xi(s) = \xi(t) + \int_{t}^{s} a(u, \xi(u)) \, du + \int_{t}^{s} \sigma(u, \xi(u)) \, dw(u).$$

Le processus $\xi_{t,\xi(t)}(s)$ sera également solution de cette équation.

Donc $\xi(s) = \xi_{t,\xi(t)}(s)$ presque sûrement. Prouvons maintenant que

$$\mathsf{P}\{\xi(s) \in A \mid \xi(t)\} = \mathsf{P}\{\xi(s) \in A \mid \mathfrak{F}_t\}.$$

Il suffit de montrer que

$$\mathsf{E}\,\zeta\lambda(\xi(s)) = \mathsf{E}\zeta\mathsf{E}(\lambda(\xi(s)) \mid \xi(t)), \tag{7}$$

où ζ est une quantité bornée $\mathfrak{F}_t$-mesurable et $\lambda(x)$ une fonction continue bornée.

Soit $\varphi(x, \omega) = \lambda(\xi_{t,x}(s))$. On a $\lambda(\xi(s)) = \varphi(\xi(t), \omega)$. Supposons d'abord que $\varphi(x, \omega)$ est de la forme

$$\varphi(x, \omega) = \sum_{k=1}^{n} \varphi_k(x)\psi_k(\omega). \tag{8}$$

Puisque $\psi_k(\omega)$ ne dépend pas de $\mathfrak{F}_t$, on aura alors

$$\mathsf{E}\zeta \sum_{k=1}^{n} \varphi_k(\xi(t))\psi_k(\omega) = \sum_{k=1}^{n} \mathsf{E}\zeta\varphi_k(\xi(t))\,\mathsf{E}\psi_k(\omega) =$$

$$= \mathsf{E}\sum_{k=1}^{n} \zeta\varphi_k(\xi(t))\,\mathsf{E}\psi_k(\omega),$$

$$\mathsf{E}\Big(\sum_{k=1}^{n} \varphi_k(\xi(t))\psi_k(\omega) \,|\, \xi(t)\Big) = \sum_{k=1}^{n} \varphi_k(\xi(t))\,\mathsf{E}\psi_k(\omega).$$

Donc (7) vaut pour le cas où $\varphi(x, \omega)$ est de la forme (8). Nous avons établi de surcroît que dans ce cas

$$\mathsf{E}(\lambda(\xi(s)) \,|\, \mathfrak{F}_t) = g(\xi(t)), \tag{9}$$

où $g(x) = \mathsf{E}\lambda(\xi_{t,x}(s))$.

On remarque que par un passage à la limite trivial, (7) et (9) s'étendent à toutes les fonctions $\varphi(x, \omega)$. Par suite

$$\mathsf{P}\{\xi(s) \in A \,|\, \mathfrak{F}_t\} = \mathsf{P}_{t,\xi(t)}(s, A),$$

où $\mathsf{P}_{t,x}(s, A) = \mathsf{P}\{\xi_{t,x}(s) \in A\}$. ∎

Montrons que moyennant des conditions auxiliaires le processus $\xi(t)$ sera de diffusion, de coefficient de diffusion $\sigma^2(t, x)$ et de coefficient de transfert $a(t, x)$. Pour cela il nous faut préalablement établir la proposition suivante.

LEMME 2. *Si $\xi_{t,x}(s)$ est une solution de l'équation* (6) *dont les solutions $a(t, x)$ et $\sigma(t, x)$ vérifient les hypothèses du théorème* 1, *alors il existe une constante H telle que*

$$\mathsf{E}\,|\xi_{t,x}(s) - x|^4 \leqslant H(s-t)^2(1 + x^4).$$

Démonstration. Soit

$$\chi_N(s) = \begin{cases} 1 \text{ pour } \sup\limits_{t \leqslant u \leqslant s} |\xi_{t,x}(u) - x| \leqslant N, \\ 0 \text{ dans le cas contraire.} \end{cases}$$

On a alors

$$(\xi_{t,x}(s) - x)\chi_N(s) =$$

$$= \chi_N(s)\Big[\int_t^s \chi_N(u)\,a(u, \xi_{t,x}(u))\,du + \int_t^s \chi_N(u)\,\sigma(u, \xi_{t,x}(u))\,dw(u)\Big].$$

L'inégalité $(a+b)^4 \leqslant 8a^4 + 8b^4$, l'inégalité de Cauchy et la propriété VII, § 1, donnent

$$\mathsf{E}\,(\xi_{t,x}(s)-x)^4\chi_N(s) \leqslant 8\mathsf{E}\left[\int_t^s \chi_N(u)\,a(u,\xi_{t,x}(u))\,du\right]^4 +$$

$$+8\mathsf{E}\left[\int_t^s \chi_N(u)\,\sigma(u,\xi_{t,x}(u))\,dw(u)\right]^4 \leqslant$$

$$\leqslant 8(s-t)^3\int_t^s \mathsf{E}\chi_N(u)\,[a(u,\xi_{t,x}(u))]^4\,du+$$

$$+8\cdot 36\,(s-t)\int_t^s \mathsf{E}\chi_N(u)\,[\sigma(u,\xi_{t,x}(u))]^4\,du \leqslant 64\,(s-t)^3\int_t^s a^4(u,x)\,du+$$

$$+64\,(s-t)^3\int_t^s \mathsf{E}\chi_N(u)\,[a(u,\xi_{t,x}(u))-a(u,x)]^4\,du+$$

$$+64\cdot 36\,(s-t)\int_t^s \sigma^4(u,x)\,du+$$

$$+64\cdot 36\,(s-t)\int_t^s \mathsf{E}\chi_N(u)\,[\sigma(u,\xi_{t,x}(u))-\sigma(u,x)]^4\,du \leqslant$$

$$\leqslant L\,(s-t)\int_t^s \mathsf{E}\chi_N(u)\,|\,\xi_{t,x}(u)-x\,|^4\,du+R\,(s-t)^2,$$

où L ne dépend que de K et $R = H_1(1 + x^4)$, où H_1 ne dépend que de K. Si donc

$$\alpha(s) = \mathsf{E}\,|\,\xi_{t,x}(s)-x\,|^4\,\chi_N(s),$$

alors

$$\alpha(s) \leqslant L\,(s-t)\int_t^s \alpha(u)\,du + R\,(s-t)^2.$$

Pour $t \leqslant s \leqslant t+1$

$$\alpha(s) \leqslant L\int_t^s \alpha(u)\,du + R\,(s-t)^2.$$

Par suite, en vertu du lemme 1,

$$\alpha(s) \leqslant R(s-t)^2 + L\int_t^s e^{L(s-u)} R(s-u)^2\,du.$$

Donc, pour $t \leqslant s < t+1$ on peut exhiber H tel que

$$\alpha(s) \leqslant H(1+x^4)(s-t)^2,$$

$$\mathsf{E}\,|\xi_{t,x}(s)-x|^4 \chi_N(s) \leqslant H(1+x^4)(s-t)^2.$$

En passant à la limite pour $N \to \infty$, on obtient ce qu'on voulait. ■

Corollaire 1. *Si $\xi(t)$ est solution de l'équation* (2), *$a(t, x)$ et $\sigma(t, x)$ vérifient les hypothèses du théorème* 1 *et* $\mathsf{E}\,|\xi(t_0)|^4 < \infty$, *alors*

$$\sup_{t_0 \leqslant t \leqslant T} \mathsf{E}\,|\xi(t)|^4 < \infty.$$

En effet, le lemme 2 implique

$$\mathsf{E}(|\xi(t)-\xi(t_0)|^4 \mid \xi(t_0)) \leqslant H(T-t_0)^2(1+|\xi(t_0)|^4),$$

$$\mathsf{E}\,|\xi(t)|^4 \leqslant 8\mathsf{E}\,|\xi(t_0)|^4 + 8\mathsf{E}\mathsf{E}(|\xi(t)-\xi(t_0)|^4 \mid \xi(t_0)).$$

Corollaire 2. *Dans les hypothèses du corollaire précédent, il existe une constante H_1 telle que*

$$\mathsf{E}\,|\xi(t)-\xi(s)|^4 \leqslant H_1(s-t)^2.$$

En effet, pour $t < s$

$$\mathsf{E}\,|\xi(s)-\xi(t)|^4 \leqslant H(s-t)^2(\mathsf{E}\,|\xi(t)|^4+1).$$

Théorème 3. *Si sont remplies les conditions du théorème* 1 *et de plus $a(t, x)$ et $\sigma(t, x)$ sont continues en t pour $t \in [t_0, T]$, alors $\sigma^2(t, x)$ et $a(t, x)$ seront respectivement coefficient de diffusion et coefficient de transfert d'un processus $\xi(t)$, solution de l'équation* (2).

Démonstration. La probabilité de passage $\mathsf{P}(t, x, s, A)$ du processus $\xi(t)$ étant confondue, en vertu du théorème 2, avec la répartition de $\xi_{t,x}(s)$, il suit du lemme 2 que

$$\int (y-x)^4 \mathsf{P}(t, x, s, dy) = \mathsf{E}\,|\xi_{t,x}(s)-x|^4 = o(s-t).$$

La remarque du théorème 6, § 4, chapitre I, nous indique que pour démontrer ce théorème il suffit de prouver que

$$\int (y-x)\,\mathsf{P}(t, x, s, dy) = \mathsf{E}\xi_{t,x}(s) - x = a(t, x)(s-t) + o(s-t) \tag{10}$$

et

$$\int (y-x)^2 \mathsf{P}(t, x, s, dy) = \mathsf{E}(\xi_{t,x}(s)-x)^2 = \sigma^2(t, x)(s-t) + o(s-t). \tag{11}$$

Ces formules se démontrent de façon analogue. Prouvons l'une d'elles, par exemple (11). De (6) il suit

$$\mathsf{E}(\xi_{t,x}(s)-x)^2 = \mathsf{E}\left(\int_t^s a(u, \xi_{t,x}(u))\,du + \int_t^s \sigma(u, \xi_{t,x}(u))\,dw(u)\right)^2.$$

Comme

$$\mathsf{E}\left[\int_t^s a(u, \xi_{t,x}(u))\,du\right]^2 \leqslant (s-t)\int_t^s \mathsf{E}a^2(u, \xi_{t,x}(u))\,du = o(s-t),$$

et

$$\mathsf{E}\left(\int_t^s \sigma(u, \xi_{t,x}(u))\,dw(u)\right)^2 = \int_t^s \mathsf{E}[\sigma(u, \xi_{t,x}(u))]^2\,du \leqslant$$

$$\leqslant K^2 \int_t^s \mathsf{E}(1+|\xi_{t,x}(u)|^2)\,du = O(s-t),$$

il vient

$$\mathsf{E}(\xi_{t,x}(s)-x)^2 = \mathsf{E}\left(\int_t^s \sigma(u, \xi_{t,x}(u))\,dw(u)\right)^2 + o(s-t) =$$

$$= \int_t^s \mathsf{E}\sigma^2(u, \xi_{t,x}(u))\,du + o(s-t).$$

L'égalité

$$\lim_{u \downarrow t} \mathsf{E}\sigma^2(u, \xi_{t,x}(u)) = \sigma(t, x),$$

qui est une conséquence du passage à la limite sous le signe de l'espérance mathématique (en vertu de l'inégalité $\sigma^4(u, \xi_{t,x}(u)) \leqslant \leqslant K^4(1+|\xi_{t,x}(u)|^2)^2$ et du corollaire 1 du théorème 2) et la continuité de $\sigma(t, x)$ donnent

$$\int_t^s \mathsf{E}\sigma^2(u, \xi_{t,x}(u))\,du = \sigma^2(t, x)(s-t) + o(s-t),$$

d'où résulte (11). ■

Nous allons construire un processus de diffusion à plusieurs dimensions $\xi(t)$ de vecteur de transfert $a(t, x)$ et d'opérateur de diffusion $B(t, x)$. Posons

$$b_k(t, x) = \sqrt{\lambda_k(t, x)}\, e_k(t, x),$$

où $e_k(t, x)$ sont les vecteurs propres, et $\lambda_k(t, x)$ les valeurs propres correspondantes de l'opérateur $B(t, x)$.

Nous chercherons le processus $\xi(t)$ comme solution de l'équation

$$d\xi(t) = a(t, \xi(t))\, dt + \sum_{k=1}^{m} b_k(t, \xi(t))\, dw_k(t), \tag{12}$$

où $w_k(t)$, $k = 1, 2, \ldots, m$, sont des processus wienériens indépendants.

Les fonctions $a(t, x)$ et $b_k(t, x)$ sont définies pour $t \in [t_0, T]$, $x \in \mathcal{R}^m$, et sont à valeurs dans $\mathcal{R}^m$. L'équation (12) équivaut à l'équation

$$\xi(t) = \xi(t_0) + \int_{t_0}^{t} a(s, \xi(s))\, ds + \sum_{k=1}^{m} \int_{t_0}^{t} b_k(s, \xi(s))\, dw_k(s). \tag{13}$$

Cette équation admet une solution pour $\xi(t_0)$: dans toute la suite on supposera que $\xi(t_0)$ ne dépend pas des processus $w_k(t)$.

Soit $\mathfrak{F}_t$ la plus petite tribu engendrée par $\xi(t_0)$ et $w_k(s) - w_k(t_0)$ pour $k = 1, \ldots, m$; $s \in [t_0, t]$. On admettra que la solution de (13) est un processus $\xi(t)$ tel qu'existent les intégrales du second membre de (13), et (13) soit presque sûrement réalisée pour $t \in [t_0, T]$.

Enonçons sous forme de théorème les principales propriétés des solutions de (13).

THÉORÈME 4. *Soient $a(t, x)$, $b_1(t, x)$, ..., $b_m(t, x)$ fonctions boréliennes définies pour $t \in [t_0, T]$, $x \in \mathcal{R}^m$, à valeurs dans $\mathcal{R}^m$. Si existe K tel que*

$$|a(t, x)|^2 + \sum_{k=1}^{m} |b_k(t, x)|^2 \leqslant K^2(1 + |x|^2),$$

$$|a(t, x) - a(t, y)| + \sum_{k=1}^{m} |b_k(t, x) - b_k(t, y)| \leqslant K|x - y|$$

$\forall x$ et $\forall y \in \mathcal{R}^m$, l'équation (13) possède une solution unique $\xi(t)$ à l'équivalence stochastique près presque sûrement continue. Cette solution $\xi(t)$ est un processus markovien dont les probabilités de passage $\mathsf{P}(t, x, s, A)$, $t < s$, sont définies par

$$\mathsf{P}(t, x, s, A) = \mathsf{P}\{\xi_{t,x}(s) \in A\},$$

où $\xi_{t,x}(s)$ *est solution de l'équation*

$$\xi_{t,x}(s)=x+\int_t^s a(u,\xi_{t,x}(u))\,du+\sum_{k=1}^m\int_t^s b_k(u,\xi_{t,x}(u))\,dw_k(u). \quad (14)$$

Si les fonctions $a(t, x)$ *et* $b_k(t, x)$ *sont continues en* t, *le processus* $\xi(t)$ *sera de diffusion, de vecteur de transfert* $a(t, x)$ *et d'opérateur de diffusion* $B(t, x)$ *tel que*

$$(B(t,x)z,z)=\sum_{k=1}^m(b_k(t,x)z)^2.$$

La démonstration de ces faits en principe ne diffère en rien de celles des théorèmes 1, 2, 3 pour les processus à une dimension.

REMARQUE. 1. *Si l'équation* (14) *dont les coefficients* $a(s, x)$ *et* $b_k(s, x)$ *remplissent les conditions du théorème* 4 *admet une solution* $\xi_{t,x}(s)$, *alors existe une constante* H *telle que*

$$\mathsf{E}\,|\xi_{t,x}(s)-x|^4\leqslant H(s-t)^2(1+|x|^4),\qquad s>t.$$

Cette proposition est équivalente à celle démontrée au lemme 2 pour un processus à une dimension.

REMARQUE. 2. *Si les coefficients* $a(t, x)$ *et* $b_k(t, x)$ *de l'équation* (12) *ne dépendent pas de* t, *c'est-à-dire l'équation est de la forme*

$$d\xi(t)=a(\xi(t))\,dt+\sum_{k=1}^m b_k(\xi(t))\,dw_k(t), \quad (15)$$

et $a(x)$, $b_k(x)$, $k = 1, \ldots, m$, *vérifient les conditions du théorème* 4, *alors la solution* $\xi(t)$ *de cette équation sera processus markovien homogène, c'est-à-dire la probabilité de passage* $\mathsf{P}(t, x, t+h, A)$ *ne dépendra pas de* t.

En effet, $\mathsf{P}(t, x, t+h, A)$ est confondue avec la répartition de $\zeta_{t,x}(h) = \xi_{t,x}(t+h)$; or le théorème 4 nous dit que $\zeta_{t,x}(h)$ sera solution de l'équation

$$d\zeta_{t,x}(h)=a(\zeta_{t,x}(h))\,dh+\sum_{k=1}^m b_k(\zeta_{t,x}(h))\,d_h[w_k(t+h)-w_k(t)]$$

avec la condition initiale $\zeta_{t,x}(0) = x$.

Comme la répartition conjointe de $[w_k(t+h) - w_k(t)]$, $k = 1, 2, \ldots, m$, ne dépend pas de t, il en sera de même de la répartition de $\zeta_{t,x}(h)$.

§ 3. Dérivabilité des solutions des équations stochastiques par rapport aux conditions initiales

Dans ce paragraphe nous allons montrer que la fonction $\xi_{t,x}(s)$, qui a été introduite au paragraphe précédent comme solution de l'équation (14), est une fonction de x, dérivable si les coefficients $a(t, x)$ et $b_k(t, x)$ le sont suffisamment de fois. Cette fonction étant

aléatoire, il convient d'abord de préciser ce qu'on entend par dérivée de $\xi_{t,x}(s)$. Pour cela il est plus commode de considérer la dérivabilité en moyenne quadratique des fonctions aléatoires. Si $\varphi(x, \omega)$ est une fonction aléatoire de $x = (x^1, \ldots, x^m) \in \mathscr{R}^m$, par $\frac{\partial \varphi}{\partial x^i}$ on entendra une variable aléatoire telle que

$$\lim_{\Delta x^i \to 0} \mathsf{E} \left| \frac{1}{\Delta x^i} [\varphi(x^1, \ldots, x^i + \Delta x^i, \ldots, x^m, \omega) - \right.$$
$$\left. - \varphi(x^1, \ldots, x^i, \ldots, x^m, \omega)] - \frac{\partial \varphi(x, \omega)}{\partial x^i} \right|^2 = 0.$$

De même que dans le paragraphe précédent, les démonstrations seront intégralement conduites pour les processus à une dimension.

THÉORÈME 1. *Si $a(t, x)$ et $\sigma(t, x)$, fonctions définies et continues pour $x \in \mathscr{R}^1$, $t \in [t_0, T]$, possèdent des dérivées partielles bornées continues $a'_x(t, x)$, $a''_{xx}(t, x)$, $\sigma'_x(t, x)$, $\sigma''_{xx}(t, x)$, alors la solution $\xi_{t,x}(s)$ de l'équation* (6), § 2, *est deux fois dérivable par rapport à x et de plus ses dérivées sont continues en x en moyenne quadratique.*

La démonstration de ce théorème s'appuie sur le lemme suivant:

LEMME 1. *Supposons que des processus $\zeta_n(t)$, $n = 0, 1, \ldots$, sont solutions des équations stochastiques*

$$\zeta_n(t) = \varphi_n(t) + \int_{t_0}^{t} \psi_n(s) \zeta_n(s)\, ds + \int_{t_0}^{t} \chi_n(s) \zeta_n(s)\, dw(s). \qquad (1)$$

Les fonctions $\varphi_n(t)$, $\psi_n(t)$ et $\chi_n(t)$ sont $\mathfrak{F}_t$-mesurables, $\forall t$, $\sup_t \mathsf{E} |\varphi_n(t)|^2 < \infty$, et existe K tel que presque sûrement $|\psi_n(s)| \leqslant K$, $|\chi_n(s)| \leqslant K$. Si $\sup_t \mathsf{E} |\varphi_n(t) - \varphi_0(t)|^2 \to 0$ pour $n \to \infty$ et pour tout t $\psi_n(t) \to \psi_0(t)$, $\chi_n(t) \to \chi_0(t)$ stochastiquement, alors $\sup_t \mathsf{E} |\zeta_n(t) - \zeta_0(t)|^2 \to 0$ pour $n \to \infty$.

D é m o n s t r a t i o n. A noter que l'existence et l'unicité de la solution de (1) ainsi que $\sup_t \mathsf{E} |\zeta_n(t)|^2 < \infty$ se démontrent exactement comme dans le théorème 1, § 2. Donc

$$\mathsf{E} |\zeta_n(t) - \zeta_0(t)|^2 \leqslant 3\mathsf{E} |\varphi_n(t) - \varphi_0(t)|^2 +$$
$$+ 3\mathsf{E} \left(\int_{t_0}^{t} [\psi_n(s) \zeta_n(s) - \psi_0(s) \zeta_0(s)]\, ds \right)^2 +$$
$$+ 3\mathsf{E} \left(\int_{t_0}^{t} (\chi_n(s) \zeta_n(s) - \chi_0(s) \zeta_0(s))\, dw(s) \right)^2 =$$

$$= 3\mathsf{E}\,|\varphi_n(t)-\varphi_0(t)|^2+3\mathsf{E}\left(\int_{t_0}^{t}\psi_n(s)\,[\zeta_n(s)-\zeta_0(s)]\,ds+\right.$$

$$\left.+\int_{t_0}^{t}\zeta_0(s)\,[\psi_n(s)-\psi_0(s)]\,ds\right)^2+3\mathsf{E}\left(\int_{t_0}^{t}\chi_n(s)\,[\zeta_n(s)-\zeta_0(s)]\,dw(s)+\right.$$

$$\left.+\int_{t_0}^{t}\zeta_0(s)\,[\chi_n(s)-\chi_0(s)]\,dw(s)\right)^2\leqslant$$

$$\leqslant \delta_n(t)+6\mathsf{E}\left(\int_{t_0}^{t}K\,|\zeta_n(s)-\zeta_0(s)|\,ds\right)^2+$$

$$+6K^2\int_{t_0}^{t}\mathsf{E}\,|\zeta_n(s)-\zeta_0(s)|^2\,ds,$$

où

$$\delta_n(t)=3\mathsf{E}\,|\varphi_n(t)-\varphi_0(t)|^2+$$

$$+6(t-t_0)\int_{t_0}^{t}\mathsf{E}\,|\zeta_0(s)|^2\,|\psi_n(s)-\psi_0(s)|^2\,ds+$$

$$+6\int_{t_0}^{t}\mathsf{E}\,|\zeta_0(s)|^2\,|\chi_n(s)-\chi_0(s)|^2\,ds.$$

Comme

$$\sup_t \delta_n(t)\leqslant 3\sup_t \mathsf{E}\,|\varphi_n(t)-\varphi_0(t)|^2+$$

$$+6(T-t_0+1)\int_{t_0}^{T}\mathsf{E}\,|\zeta_0(s)|^2\,(|\psi_n(s)-\psi_0(s)|^2+|\chi_n(s)-\chi_0(s)|^2)\,ds,$$

$\delta_n(t)$ tend uniformément en t vers zéro (l'intégrale tend vers 0 d'après le théorème de Lebesgue, car l'intégrande est une fonction majorée par $4K^2\,|\zeta_0(s)|^2$ et tend stochastiquement vers 0).

Si l'on pose $H=6K^2(T-t_0+1)$, on obtient

$$\mathsf{E}\,|\zeta_n(t)-\zeta_0(t)|^2\leqslant \sup_t \delta_n(t)+H\int_{t_0}^{t}\mathsf{E}\,|\zeta_n(s)-\zeta_0(s)|^2\,ds.$$

En vertu du lemme 1, § 2, il vient

$$\mathsf{E}\,|\zeta_n(t)-\zeta_0(t)|^2\leqslant \sup_t \delta_n(t)\,e^{H(T-t_0)}.\ ■$$

REMARQUE. *Le lemme est vrai pour des processus dépendant d'un paramètre α continu, pourvu que les limites respectives existent pour $\alpha \to 0$.*

Voyons maintenant la démonstration du théorème. Posons

$$\zeta_{x,\,\Delta x}(s) = \frac{1}{\Delta x}\,[\xi_{t,\,x+\Delta x}(s) - \xi_{t,\,x}].$$

Le processus $\zeta_{x,\,\Delta x}(s)$ sera alors solution de l'équation

$$\zeta_{x,\,\Delta x}(s) = 1 + \int_t^s \psi_{x,\,\Delta x}(u)\,\zeta_{x,\,\Delta x}(u)\,du + \int_t^s \chi_{x,\,\Delta x}(u)\,\zeta_{x,\,\Delta x}(u)\,dw(u),$$

où

$$\psi_{x,\,\Delta x}(s) = \frac{a(s,\,\xi_{t,\,x+\Delta x}(s)) - a(s,\,\xi_{t,\,x}(s))}{\xi_{t,\,x+\Delta x}(s) - \xi_{t,\,x}(s)},$$

$$\chi_{x,\,\Delta x}(s) = \frac{\sigma(s,\,\xi_{t,\,x+\Delta x}(s)) - \sigma(s,\,\xi_{t,\,x}(s))}{\xi_{t,\,x+\Delta x}(s) - \xi_{t,\,x}(s)}.$$

$a(s,\,x)$ et $\sigma(s,\,x)$ possédant des dérivées par rapport à x continues, il existe une constante K telle que presque sûrement $|\psi_{x,\,\Delta x}(s)| \leqslant K$ et $|\chi_{x,\,\Delta x}(s)| \leqslant K$. Donc

$$\mathsf{E}|\zeta_{x,\,\Delta x}(s)|^2 \leqslant 3 + 3(T - t_0)\int_t^s K^2 \mathsf{E}|\zeta_{x,\,\Delta x}(u)|^2\,du +$$
$$+ 3K^2 \int_t^s \mathsf{E}|\zeta_{x,\,\Delta x}(u)|^2\,du \leqslant 3 + H\int_t^s \mathsf{E}|\zeta_{x,\,\Delta x}(u)|^2\,du,$$

où $H = 3(T - t_0 + 1)K^2$. Du lemme 1, § 2, il suit

$$\mathsf{E}|\zeta_{x,\,\Delta x}(s)|^2 \leqslant 3e^{H(T-t_0)},$$

et par suite pour un H_1

$$\mathsf{E}|\xi_{t,\,x+\Delta x}(s) - \xi_{t,\,x}(s)|^2 \leqslant H_1(\Delta x)^2. \tag{2}$$

La dernière relation exprime que $\xi_{t,\,x+\Delta x}(s) - \xi_{t,\,x}(s) \to 0$ stochastiquement avec Δx, donc

$$\psi_{x,\,\Delta x}(s) \to a'_x(s,\,\xi_{t,\,x}(s)),\quad \chi_{x,\,\Delta x}(s) \to \sigma'_x(s,\,\xi_{t,\,x}(s))$$

stochastiquement pour $\Delta x \to 0$. Soit $\zeta_x(s)$ solution de l'équation

$$\zeta_x(s) = 1 + \int_t^s a'_x(u,\,\xi_{t,\,x}(u))\,\zeta_x(u)\,du +$$
$$+ \int_t^s \sigma'_x(u,\,\xi_{t,\,x}(u))\,\zeta_x(u)\,dw(u). \tag{3}$$

Du lemme 1 il résulte que $\mathsf{E}|\zeta_{x,\,\Delta x}(s)-\zeta_x(s)|^2\to 0$ avec Δx, i.e.

$$\zeta_x(s)=\frac{\partial}{\partial x}\xi_{t,\,x}(s).$$

A remarquer que la relation (3) entraîne

$$\zeta_x(s)=\exp\left\{\int_t^s \left(a'_x(u,\ \xi_{t,\,x}(u))-\frac{1}{2}[\sigma'_x(u,\ \xi_{t,\,x}(u))]^2\right)du+\right.$$
$$\left.+\int_t^s \sigma'_x(u,\ \xi_{t,\,x}(u))\,dw(u)\right\}. \qquad (4)$$

En effet, supposons que le processus $\zeta(s)$ est tel que $\zeta(t)=1$ et

$$d\zeta(s)=a(s)\,\zeta(s)\,ds+\sigma(s)\,\zeta(s)\,dw(s),$$

où $a(s)$ et $b(s)$ sont des fonctions bornées. La formule de Ito donne

$$d\left[\zeta(s)\exp\left\{-\int_t^s\left[a(u)-\frac{1}{2}\sigma^2(u)\right]du-\int_t^s\sigma(u)\,dw(u)\right\}\right]=$$
$$=[d\zeta(s)]\exp\left\{-\int_t^s\left[a(u)-\frac{1}{2}\sigma^2(u)\right]du-\int_t^s\sigma(u)\,dw(u)\right\}+$$
$$+\zeta(s)\exp\left\{-\int_t^s\left[a(u)-\frac{1}{2}\sigma^2(u)\right]du-\int_t^s\sigma(u)\,dw(u)\right\}\times$$
$$\times\left[\left(-a(s)+\frac{1}{2}\sigma^2(s)+\frac{1}{2}\sigma^2(s)\right)ds-\sigma(s)\,dw(s)\right]-$$
$$-\zeta(s)\exp\left\{-\int_t^s\left[a(u)-\frac{1}{2}\sigma^2(u)\right]du-\int_t^s\sigma(u)\,dw(u)\right\}\sigma^2(s)\,ds=0,$$

donc

$$\zeta(s)\exp\left\{-\int_t^s\left[a(u)-\frac{1}{2}\sigma^2(u)\right]du-\int_t^s\sigma(u)\,dw(u)\right\}=c.$$

En posant $s=t$ on trouve $c=1$. Ce qui prouve la formule (4). Dans la suite on aura besoin du

LEMME 2. *Si* $|\varphi(u)|\leqslant N$, $u\in[s,\ t]$, *alors*

$$\mathsf{E}\exp\left\{\int_s^t\varphi(u)\,dw(u)\right\}\leqslant\exp\left\{\frac{1}{2}N^2(s-t)\right\}. \qquad (5)$$

Démonstration. Si $\varphi(u)$ est une fonction en escalier: $\varphi(u)=\varphi(t_k)$ pour $u\in[t_k,\ t_{k+1}]$, où $t=t_0<t_1<\ldots<t_r=s$,

alors on obtient (5) grâce à l'inégalité

$$\mathsf{E}(\exp\{\varphi(t_k)[w(t_{k+1})-w(t_k)]\}|\mathfrak{F}_{t_k})=$$
$$=\exp\left\{\frac{1}{2}\varphi^2(t_k)(t_{k+1}-t_k)\right\}\leqslant\exp\left\{\frac{1}{2}N(t_{k+1}-t_k)\right\}.$$

Dans le cas général on établit (5) par passage à la limite. ■

Le lemme 2, le fait que $a_x^{\bullet}$ et $\sigma_x^{\bullet}$ sont bornées et (4) entraînent que

$$\mathsf{E}(\zeta_x(s))^m \ \forall m>0$$

est uniformément bornée. De la formule (4) on déduit aisément l'existence de $\frac{\partial}{\partial x}\zeta_x(s)$ em moyenne quadratique et

$$\frac{\partial^2}{\partial x^2}\xi_{t,x}(s)=\frac{\partial}{\partial x}\zeta_x(s)=\exp\left\{\int_t^s\left(a_x'(u,\ \xi_{t,x}(u))-\right.\right.$$
$$\left.\left.-\frac{1}{2}[\sigma_x'(u,\ \xi_{t,x}(u))]^2\right)du+\int_t^s\sigma_x'(u,\ \xi_{t,x}(u))\,dw(u)\right\}\times$$
$$\times\left[\int_t^s[a_{xx}''(u,\ \xi_{t,x}(u))-\sigma_x'(u,\ \xi_{t,x}(u))\,\sigma_{xx}''(u,\ \xi_{t,x}(u))]\,\zeta_x(u)\,du+\right.$$
$$\left.+\int_t^s\sigma_{xx}''(u,\ \xi_{t,x}(u))\,\zeta_x(u)\,dw(u)\right].\qquad(6)$$

Cette relation entraîne la continuité en moyenne quadratique de $\frac{\partial^2}{\partial x^2}\xi_{t,x}(s)$. ■

Pour les processus à plusieurs dimensions on a le

THÉORÈME 2. *Si $\xi_{t,x}(s)$ est solution de l'équation (14), § 2, et que $a(t,x)$, $b_1(t,x)$, ..., $b_m(t,x)$, fonctions définies et continues pour $t\in[t_0, T]$, $x\in\mathscr{R}^m$, possèdent des dérivées premières et secondes continues bornées par rapport à toutes les variables x^1, ..., x^m, alors la fonction $\xi_{t,x}(s)$ est deux fois continûment dérivable en moyenne quadratique par rapport à x et de plus les dérivées*

$$\frac{\partial}{\partial x^i}\xi_{t,x}(s),\quad \frac{\partial^2}{\partial x^i\,\partial x^j}\xi_{t,x}(s),$$

comme fonctions de x, seront continues en moyenne quadratique.

La démonstration de ce théorème sera omise, car elle suit le même schéma que celle du théorème 1.

REMARQUE 1. *Si les conditions du théorème 2 sont réunies et que $f(x)$ soit une fonction continue bornée possédant des dérivées pre-*

mières et secondes continues et bornées, alors $u(x) = \mathsf{E}f(\xi_{t,\,x}(s))$ *est deux fois continûment dérivable par rapport à* x.

Nous allons prouver cette assertion pour les processus à une dimension. Montrons que

$$u'_x(x) = \mathsf{E}f'_x(\xi_{t,\,x}(s))\,\zeta_x(s). \tag{7}$$

En effet,

$$\mathsf{E}\left[\frac{f(\xi_{t,\,x+\Delta x}(s)) - f(\xi_{t,\,x}(s))}{\Delta x} - f'_x(\xi_{t,\,x}(s))\,\zeta_x(s)\right]^2 \leqslant$$

$$\leqslant 2\mathsf{E}\left[\frac{f(\xi_{t,\,x+\Delta x}(s)) - f(\xi_{t,\,x}(s))}{\xi_{t,\,x+\Delta x}(s) - \xi_{t,\,x}(s)}\,[\zeta_{x,\,\Delta x}(s) - \zeta_x(s)]\right]^2 +$$

$$+ 2\mathsf{E}\left[\frac{f(\xi_{t,\,x+\Delta x}(s)) - f(\xi_{t,\,x}(s))}{\xi_{t,\,x+\Delta x}(s) - \xi_{t,\,x}(s)} - f'_x(\xi_{t,\,x}(s))\right]^2 [\zeta_x(s)]^2 \to 0,$$

puisque $\dfrac{f(x) - f(y)}{x - y}$ est bornée,

$$\frac{f(\xi_{t,\,x+\Delta x}(s)) - f(\xi_{t,\,x}(s))}{\xi_{t,\,x+\Delta x}(s) - \xi_{t,\,x}(s)} - f'_x(\xi_{t,\,x}(s))$$

stochastiquement convergente vers zéro et

$$\lim_{\Delta x \to 0} \mathsf{E}\,|\zeta_{x,\,\Delta x}(s) - \zeta_x(s)|^2 = 0, \quad \mathsf{E}\,|\zeta_x(s)|^2 < \infty.$$

Or

$$\left|\frac{u(x + \Delta x) - u(x)}{\Delta x} - \mathsf{E}f'_x(\xi_{t,\,x}(s))\,\zeta_x(s)\right| \leqslant$$

$$\leqslant \left\{\mathsf{E}\left[\frac{1}{\Delta x}\left(f(\xi_{t,\,x+\Delta x}(s)) - f(\xi_{t,\,x}(s))\right) - f'_x(\xi_{t,\,x}(s))\,\zeta_x(s)\right]^2\right\}^{1/2}$$

D'où (7). De façon analogue,

$$u''_{xx} = \mathsf{E}f''_{xx}(\xi_{t,\,x}(s))\,\zeta_x^2(s) + \mathsf{E}f'_x(\xi_{t,\,x}(s))\,\frac{\partial}{\partial x}\,\zeta_x(s). \tag{8}$$

La continuité de u'_x et u''_{xx} découle de la continuité en moyenne quadratique des processus $\zeta_x(s)$ et $\dfrac{\partial}{\partial x}\,\zeta_x(s)$.

Remarque 2. *Les processus*

$$\xi_{t,\,x}(s), \quad \frac{\partial}{\partial x}\,\xi_{t,\,x}(s), \quad \frac{\partial^2}{\partial x^2}\,\xi_{t,\,x}(s)$$

sont des fonctions de t *stochastiquement continues à* s *fixe*, $t_0 \leqslant s \leqslant T$, *uniformément en* x *sur tout compact.*

Soit $t < t' < s$; il vient

$$\xi_{t,\,x}(s) - \xi_{t',\,x}(s) = \xi_{t,\,x}(t') - x + \int_{t'}^{s} (a(u,\,\xi_{t\,x}(u)) -$$

$$-a(u, \xi_{t',x}(u)))\,du + \int_{t'}^{s} [\sigma(u, \xi_{t,x}(u)) - \sigma(u, \xi_{t',x}(u))]\,dw(u).$$

Donc pour un H on a

$$\mathsf{E}[\xi_{t,x}(s) - \xi_{t',x}(s)]^2 \leqslant 3\mathsf{E}|\xi_{t,x}(t') - x|^2 + H\int_{t'}^{s} \mathsf{E}[\xi_{t,x}(u) - \xi_{t',x}(u)]^2\,du,$$

d'où il suit

$$\mathsf{E}[\xi_{t,x}(s) - \xi_{t',x}(s)]^2 \leqslant 3\mathsf{E}|\xi_{t,x}(t') - x|^2 e^{H(s-t')}.$$

Or $\mathsf{E}\,|\,\xi_{t,x}(t') - x\,|^2 = O\,(t' - t)$ uniformément sur tout compact en vertu du lemme 2, § 2. Donc

$$\mathsf{E}|\xi_{t,x}(s) - \xi_{t',x}(s)|^2 = O(t' - t).$$

En se servant de (4) et (6) on obtient des majorations analogues pour

$$\frac{\partial}{\partial x}\xi_{t,x}(s), \quad \frac{\partial^2}{\partial x^2}\xi_{t,x}(s).$$

En regroupant les remarques 1 et 2 on obtient le théorème suivant.

THÉORÈME 3. *Si* $\xi_{t,x}(s)$ *est une solution de l'équation* (14), § 2, *dont les coefficients* $a\,(t, x)$, $b_1\,(t, x)$, . . ., $b_m\,(t, x)$ *vérifient les conditions du théorème* 2, *et* $f\,(x)$ *une fonction définie sur* $\mathcal{R}^m$, *continue, bornée et possédant des dérivées première et seconde continues et bornées, alors la fonction*

$$u(t, x) = \mathsf{E}f(\xi_{t,x}(s)),$$

définie pour $t_0 \leqslant t \leqslant s$, $x \in \mathcal{R}^m$, *est continue, possède des dérivées première et seconde par rapport à* x *continues en* t *et* x.

§ 4. Méthode des équations différentielles

Dans ce paragraphe nous allons établir des équations différentielles qui nous permettront de déterminer les répartitions de certaines fonctionnelles de processus de diffusion. Au passage on donnera une nouvelle déduction de l'équation (inverse) de Kolmogorov pour processus de diffusion.

Soit $\xi\,(t)$ solution de l'équation (13), § 2. On a vu au § 2 que la répartition conditionnelle du processus $\xi\,(s)$ sur $[t, T]$, sachant que $\xi\,(t) = x$, était confondue avec la répartition de $\xi_{t,x}\,(s)$. Appelons $\mathsf{E}_{t,x}\eta$ l'espérance mathématique conditionnelle des variables aléatoires η, fonctions de la trajectoire du processus $\xi\,(s)$ sur $[t, T]$, sachant que $\xi\,(t) = x$. De ce qui précède il suit

$$\mathsf{E}_{t,x}\eta = \mathsf{E}\hat{\eta},$$

où $\hat{\eta}$ se déduit de η par substitution de la trajectoire $\xi_{t,x}(s)$ à la trajectoire $\xi(s)$. Pour définir les probabilités de passage du processus il suffit de déterminer

$$\mathsf{E}_{t,x}\varphi(\xi(s)) = \int \varphi(y)\,\mathsf{P}(t, x, s, dy)$$

pour des fonctions suffisamment de fois dérivables.

THÉORÈME 1. *Si $\xi(t)$ est un processus remplissant les conditions du théorème 2, § 3, et $\varphi(x)$ une fonction bornée, continue et possédant des dérivées première et seconde continues et bornées, alors la fonction*

$$u(t, x) = \mathsf{E}_{t,x}\varphi(\xi(s)),\quad t \in [t_0, s[,$$

possède des dérivées première et seconde par rapport à x^i continues, est dérivable par rapport à t, vérifie l'équation

$$\frac{\partial}{\partial t}u(t, x) + \sum_{i=1}^{m} a^i(t, x)\frac{\partial}{\partial x^i}u(t, x) + \\ + \frac{1}{2}\sum_{k,i,j=1}^{m} b_k^i(t, x)\,b_k^j(t, x)\frac{\partial^2}{\partial x^i\,\partial x^j}u(t, x) = 0 \qquad (1)$$

et la condition $\lim_{t\uparrow s} u(t, x) = \varphi(x)$. Ici a^i, b_k^i, x^i sont les coordonnées des vecteurs a, b_k, x respectivement.

Démonstration. Le théorème 3, § 3, nous dit que la fonction $u(t, x)$ est dérivable par rapport à x et que ses dérivées partielles sont continues et bornées. A noter ensuite que

$$u(t, x) = \mathsf{E}_{t,x}\varphi(\xi(s)) = \mathsf{E}_{t,x}\mathsf{E}_{t+\Delta t,\,\xi(t+\Delta t)}\varphi(\xi(s)) = \\ = \mathsf{E}_{t,x}u(t+\Delta t, \xi(t+\Delta t)).$$

La formule de Ito donne

$$u(t+\Delta t, \xi(t+\Delta t)) - u(t+\Delta t, \xi(t)) = \\ = \int_t^{t+\Delta t}\Big[\sum_{j=1}^{m}\frac{\partial u}{\partial x^j}(t+\Delta t, \xi(s))\,a^j(s, \xi(s)) + \\ + \frac{1}{2}\sum_{k,i,j=1}^{m} b_k^i(s, \xi(s))\,b_k^j(s, \xi(s))\,\frac{\partial^2}{\partial x^i\,\partial x^j}u(t+\Delta t, \xi(s))\Big]ds + \\ + \int_t^{t+\Delta t}\sum_{k,i=1}^{m} b_k^i(s, \xi(s))\frac{\partial}{\partial x^i}u(t+\Delta t, \xi(s))\,dw_k(s).$$

Donc

$$\mathsf{E}_{t,x}u(t+\Delta t,\ \xi(t+\Delta t))-\mathsf{E}_{t,x}u(t+\Delta t,\ \xi(t))=$$

$$=\int_t^{t+\Delta t}\mathsf{E}_{t,x}\Big[\sum_{j=1}^m a^j(s,\ \xi(s))\frac{\partial}{\partial x^j}u(t+\Delta t,\ \xi(s))+$$

$$+\frac{1}{2}\sum_{k,i,j=1}^m b_k^i(s,\ \xi(s))\,b_k^j(s,\ \xi(s))\frac{\partial^2}{\partial x^i\,\partial x^j}u(t+\Delta t,\ \xi(s))\Big]\,ds.$$

Comme

$$\mathsf{E}_{t,x}u(t+\Delta t,\ \xi(t))=u(t+\Delta t,\ x),$$

il vient

$$u(t,\ x)-u(t+\Delta t,\ x)=\mathsf{E}_{t,x}\Big[\sum_{i=1}^m\frac{\partial}{\partial x^i}u(t+\Delta t,\ \xi(s'))\,a^i(s',\ \xi(s'))+$$

$$+\frac{1}{2}\sum_{k,i,j=1}^m b_k^i(s',\ \xi(s'))\,b_k^j(s',\ \xi(s'))\frac{\partial^2}{\partial x^i\,\partial x^j}u(t+\Delta t,\ \xi(s'))\Big]\,\Delta t,$$

où $s'\in\,]t,\ t+\Delta t[$.

Comme $\dfrac{\partial u}{\partial x^i}$, $\dfrac{\partial^2 u}{\partial x^i\,\partial x^j}$ sont bornées et $s'\to t$ pour $\Delta t\to 0$, le théorème de Lebesgue relatif au passage à la limite sous le signe de l'intégrale nous apprend que

$$\frac{u(t,\ x)-u(t+\Delta t,\ x)}{\Delta t}\to\mathsf{E}_{t,x}\Big[\sum_{i=1}^m\frac{\partial}{\partial x^i}u(t,\ \xi(t))\,a^i(t,\ \xi(t))+$$

$$+\frac{1}{2}\sum_{k,i,j=1}^m\frac{\partial^2}{\partial x^i\,\partial x^j}u(t,\ \xi(t))\,b_k^i(t,\ \xi(t))\,b_k^j(t,\ \xi(t))\Big].$$

L'équation (1) est établie. Que $\mathsf{E}_{t,x}\varphi(\xi(s))\to\varphi(x)$ découle de la relation $\mathsf{E}_{t,x}\varphi(\xi(s))=\mathsf{E}\varphi(\xi_{t,x}(s))$ et de la continuité de $\xi_{t,x}(s)$. ■

Déduisons maintenant les équations qui nous permettront de déterminer la répartition de la variable aléatoire

$$J=\int_{t_0}^T f(s,\ \xi(s))\,ds,$$

où $f(s, x)$ est une fonction suffisamment de fois dérivable, $\xi(s)$ un processus, solution de l'équation (13), § 2. Soit la fonction

$$v_\lambda(t, x) = \mathsf{E}_{t,x} \exp\left\{\lambda \int_t^T f(s, \xi(s))\, ds\right\}. \tag{2}$$

Pour déterminer la répartition de la variable J il suffit de connaître la fonction $v_\lambda(t, x)$ pour $t \in [t_0, T]$, $x \in \mathscr{R}^m$, et tous les λ imaginaires, puisque alors $v_\lambda(t_0, x)$ nous donnera la fonction caractéristique conditionnelle de J sachant que $\xi(t_0) = x$. En intégrant $v_\lambda(t_0, x)$ par rapport à la répartition initiale on obtient la fonction caractéristique conditionnelle de J.

THÉORÈME 2. *Si $\xi(t)$ remplit les conditions du théorème 2, § 3, et $f(t, x)$, $\frac{\partial}{\partial x^i} f(t, x)$, $\frac{\partial^2}{\partial x^i \partial x^j} f(t, x)$ $(i, j = 1, \ldots, m)$ sont continues et bornées, alors pour $t \in [t_0, T]$ la fonction $v_\lambda(t, x)$ vérifie l'équation*

$$\frac{\partial}{\partial t} v_\lambda(t, x) + \sum_{i=1}^m a^i(t, x) \frac{\partial}{\partial x^i} v_\lambda(t, x) +$$

$$+ \frac{1}{2} \sum_{i, j, k=1}^m b_k^i(t, x)\, b_k^j(t, x) \frac{\partial^2}{\partial x^i \partial x^j} v_\lambda(t, x) + \lambda f(t, x)\, v_\lambda(t, x) = 0 \tag{3}$$

et la condition $\lim\limits_{t \uparrow T} v_\lambda(t, x) = 1$.

Démonstration. La dernière condition est évidente. Que $v_\lambda(t, x)$ soit continue et dérivable et $\frac{\partial}{\partial x^i} v_\lambda(t, x)$, $\frac{\partial^2}{\partial x^i \partial x^j} v_\lambda(t, x)$ continues et bornées découle de la formule (2) et de la dérivabilité de $\xi_{t,x}(s)$ et $f(s, x)$ par rapport à x exactement comme dans la démonstration du théorème 3, § 3. Si dans la relation

$$\lambda \int_{t'}^{t''} \exp\left\{\lambda \int_s^T f(u, \xi(u))\, du\right\} f(s, \xi(s))\, ds =$$

$$= \exp\left\{\lambda \int_{t'}^T f(u, \xi(u))\, du\right\} - \exp\left\{\lambda \int_{t''}^T f(u, \xi(u))\, du\right\}$$

on prend $\mathsf{E}_{t',x}$ pour $t' < t''$, on obtient

$$v_\lambda(t', x) - \mathsf{E}_{t',x} \exp\left\{\lambda \int_{t''}^T f(u, \xi(u))\, du\right\} =$$

$$=\lambda \int_{t'}^{t''} \mathsf{E}_{t', x} f(s, \xi(s)) \mathsf{E}_{s, \xi(s)} \exp\left\{\lambda \int_{s}^{T} f(u, \xi(u))\, du\right\} ds.$$

Or

$$\mathsf{E}_{t', x} \exp\left\{\lambda \int_{t''}^{T} f(s, \xi(s))\, ds\right\} =$$

$$= \mathsf{E}_{t', x} \mathsf{E}_{t'', \xi(t'')} \exp\left\{\lambda \int_{t''}^{T} f(s, \xi(s))\, ds\right\} = \mathsf{E}_{t', x} v_\lambda(t'', \xi(t'')).$$

Donc

$$v_\lambda(t', x) - \mathsf{E}_{t', x} v_\lambda(t'', \xi(t'')) = \lambda \int_{t'}^{t''} \mathsf{E}_{t', x} f(s, \xi(s)) v_\lambda(s, \xi(s))\, ds.$$

Comme

$$f(s, \xi_{t', x}(s)) v_\lambda(s, \xi(s)) - f(t', x) v_\lambda(t', x) \to 0$$

stochastiquement pour $t' \to s$, existe la limite

$$\lim_{\substack{t'' \to t \\ t' \to t}} \frac{1}{t'' - t'} \int_{t'}^{t''} \mathsf{E}_{t', x} f(s, \xi(s)) v_\lambda(s, \xi(s))\, ds = f(t, x) v_\lambda(t, x).$$

Par suite, existe

$$\lim_{\substack{t'' - t' \downarrow 0 \\ t' \to t}} \frac{v_\lambda(t', x) - \mathsf{E}_{t', x} v_\lambda(t'', \xi(t''))}{t'' - t'} =$$

$$= \lim_{\substack{t'' - t' \downarrow 0 \\ t' \to t}} \left[\frac{v_\lambda(t', x) - v_\lambda(t'', x)}{t'' - t'} + \frac{v_\lambda(t'', x) - \mathsf{E}_{t', x} v_\lambda(t'', \xi(t''))}{t'' - t'}\right].$$

Or, on a vu dans la démonstration du théorème 1 que

$$\lim_{\substack{t'' - t' \downarrow 0 \\ t' \to t}} \frac{\mathsf{E}_{t', x} v_\lambda(t'', \xi(t'')) - v_\lambda(t'', x)}{t'' - t'} = \sum_{i=1}^{m} a^i(t, x) \frac{\partial}{\partial x^i} v_\lambda(t, x) +$$

$$+ \frac{1}{2} \sum_{i, j, k=1}^{m} b_k^i(t, x) b_k^j(t, x) \frac{\partial^2}{\partial x^i \partial x^j} v_\lambda(t, x).$$

Donc existe

$$\lim_{\substack{t'' - t' \downarrow 0 \\ t' \to t}} \frac{v_\lambda(t'', x) - v_\lambda(t', x)}{t'' - t'} = \frac{\partial}{\partial t} v_\lambda(t, x)$$

et l'équation (3) a bien lieu. ■

§ 5. Problèmes aux limites pour processus de diffusion

Soit $\xi(t)$ solution de l'équation différentielle stochastique (12), § 2; on admet que les coefficients de (12) remplissent les conditions du théorème 4, § 2, de sorte que $\xi(t)$ existe et est unique. Soient D domaine connexe de $[t_0, T] \times \mathcal{R}^m$, $D_t = \{x: (t; x) \in D\}$ domaine de $\mathcal{R}^m$. Posons

$$\tau = \inf\{s : s \in [t_0, T],\ \xi(s) \notin D_t\},$$

si l'ensemble figurant sous le signe inf n'est pas vide et $\tau = T$ dans le cas contraire. τ est instant markovien par rapport aux tribus $\mathfrak{F}_t$ engendrées par $\xi(t_0)$ et $w_k(s) - w_k(t_0)$, $k = 1, \ldots, m$, $s \in [t_0, t]$. En effet, si D^c est le complémentaire de D dans $[t_0, T] \times \mathcal{R}^m$ et $\rho((t; x), D^c)$ la distance du point $(t; x)$ à D^c, on a $\tau = \sup \tau_n$, où

$$\tau_n = \inf\{s: \rho((s; \xi(s)), D^c) < 1/n\},$$

si l'ensemble figurant sous le signe inf n'est pas vide, et $\tau = T$ dans le cas contraire. Que τ_n soit instant markovien résulte de

$$\{\tau_n > t\} = \bigcap_k \{\omega : \rho((s_k; \xi(s_k)), D^c) \geqslant 1/n\},$$

où $\{s_k\}$ est un ensemble partout dense dans $[t_0, t]$. L'instant τ est appelé *instant de première sortie* du domaine D. Nous allons nous intéresser aux répartitions suivantes rattachées à τ:

1) la répartition de $\xi(t)$ sachant que $\tau > t$;
2) la répartition de τ;
3) la répartition de $\xi(\tau)$.

Appelons $Lu(t, x)$ l'opérateur différentiel parabolique défini par le premier membre de la relation (1), § 4.

THÉORÈME 1. *Soient $\Phi(t, x)$ fonction définie et continue sur l'adhérence de $D \cap [t_0, t_1) \times \mathcal{R}^m$, deux fois continûment dérivable par rapport à x et dérivable par rapport à t en tous les points intérieurs de cet ensemble,* $L\Phi(t, x) = 0$ *et* $\Phi(t, x) = 0$ *pour* $t \in]t_0, t_1[$, $x \in D'_t$, *où D'_t est la frontière de D_t dans $\mathcal{R}^m$, $\Phi(t_1, x) = f(x)$. Alors*

$$\Phi(t, x) = \mathsf{E} f(\xi_{t, x}(t_1))\, \chi_{\{\tau_{t, x} \geqslant t_1\}},$$

où $\tau_{t, x}$ est l'instant de première sortie de D du processus $\xi_{t, x}$, solution de l'équation (14), § 2.

D é m o n s t r a t i o n. La formule de Ito (cf. (11), § 1) nous permet de prouver pour $\tau_{t, x} > t_1$ que

$$\Phi(t_1, \xi(t_1)) - \Phi(t, x) = \int_t^{t_1} L\Phi(s, \xi_{t, x}(s))\, ds +$$

$$+\frac{1}{2}\sum_{k=1}^{m}\int_{t}^{t_1}\sum_{i=1}^{m}\frac{\partial\Phi}{\partial x}(s,\ \xi_{t,x}(s))\,b_{ki}(s,\ \xi_{t,x}(s))\,dw_k(s). \qquad (1)$$

On établit la formule (1) de la manière suivante. Soit D_n suite de compacts simplement connexes dans $[t_0,\ T]\times\mathscr{R}^m$ tels que $D_n\subset\subset D_{n+1}$, $\bigcup D_n = D$. Désignons par $\Phi_n(t,\ x)$ une fonction confondue avec $\Phi(t, x)$, deux fois continûment dérivable par rapport à x et dérivable par rapport à t, bornée avec ses dérivées sur $[t_0,\ T]\times\mathscr{R}^m$. La fonction $\Phi_n(t,\ x)$ est alors justiciable de la formule de Ito et

$$\Phi_n(t_1,\ \xi_{t,x}(t_1))-\Phi_n(t,\ x)=\int_{t}^{t_1}L\Phi_n(s,\ \xi_{t,x}(s))\,ds+$$

$$+\sum_{k=1}^{m}\int_{t}^{t_1}\sum_{i=1}^{m}\frac{\partial}{\partial x^i}\Phi_n(s,\ \xi_{t,x}(s))\,b_{ki}(s,\ \xi_{t,x}(s))\,dw_k(s).$$

Il est évident que $t_1<\tau_{t,x}$ implique $t_1<\tau_{t,x}^{(n)}$ pour n assez grand, où $\tau_{t,x}^{(n)}$ est l'instant de première sortie de D_n du processus $\xi_{t,x}(s)$. Or, pour $s\leqslant\tau_{t,x}^{(n)}$

$$\Phi_n(s,\ \xi_{t,x}(s))=\Phi(s,\ \xi_{t,x}(s)),$$

$$L\Phi_n(s,\ \xi_{t,x}(s))=L\,\Phi(s,\ \xi_{t,x}(s)),$$

$$\frac{\partial}{\partial x^i}\Phi_n(s,\ \xi_{t,x}(s))=\frac{\partial}{\partial x^i}\Phi(s,\ \xi_{t,x}(s)).$$

Ce qui prouve la formule (1).

Soit $\tau_n'=t_1\wedge\tau_{t,x}^{(n)}$. Alors τ_n' est instant markovien et $L\Phi\times(s,\ \xi_{t,x}(s))=0$ pour $s<\tau_n'$. Donc

$$\Phi(\tau_n',\ \xi_{t,x}(\tau_n'))=$$

$$=\Phi(t,\ x)+\sum_{k=1}^{m}\int_{t}^{\tau_n'}\sum_{i=1}^{m}\frac{\partial\Phi}{\partial x_i}(s,\ \xi_{t,x}(s))\,b_{ki}(s,\ \xi_{t,x}(s))\,dw_k(s).$$

En prenant l'espérance mathématique et vu que $\frac{\partial\Phi}{\partial x^i}(s,\ \xi_{t,x}(s))$ sont bornées pour $s\leqslant\tau_n'$, car $(s,\ \xi_{t,x}(s))\in D_n$, on trouve

$$\mathsf{E}\Phi(\tau_n',\ \xi(\tau_n'))=\Phi(t,\ x).$$

En passant à la limite pour $n\to\infty$ on obtient

$$\mathsf{E}\Phi(\tau_{t,x}\wedge t_1,\ \xi(\tau_{t,x}\wedge t_1))=\Phi(t,\ x).$$

Or, $\Phi(\tau_{t,x} \wedge t_1,\ \xi(\tau_{t,x} \wedge t_1)) = \Phi(\tau_{t,x},\ \xi(\tau_{t,x})) = 0$ pour $t_1 > \tau_{t,x}$, car $\xi(\tau_{t,x}) \in D'_{\tau_{t,x}}$. Donc

$$\mathsf{E}\Phi(\tau_{t,x} \wedge t_1,\ \xi(\tau_{t,x} \wedge t_1)) = \mathsf{E}\Phi(t_1,\ \xi_{t,x}(t_1))\,\chi_{\{\tau_{t,x} \geqslant t_1\}} =$$
$$= \mathsf{E}f(\xi_{t,x}(t_1))\,\chi_{\{\tau_{t,x} \geqslant t_1\}}. \ \blacksquare$$

Pour déterminer la répartition conjointe de τ et $\xi(\tau)$ il suffit de définir $\mathsf{E}\psi(\tau,\ \xi(\tau))$ pour toutes les fonctions $\psi(t, x)$ suffisamment de fois dérivables, données sur la frontière Γ du domaine D.

Théorème 2. *Soient $v(t, x)$ fonction continue dans $D \cup \Gamma$ pour $(t;\ x) \in D$, $Lv(t, x) = 0$ et $v(t, x) = \psi(t, x)$ pour $t > t_0$, $(t;\ x) \in \Gamma$. Alors*

$$v(t,\ x) = \mathsf{E}\psi(\tau_{t,x},\ \xi_{t,x}(\tau_{t,x})).$$

Démonstration. La formule de Ito et les raisonnements du théorème 1 nous conduisent à l'égalité

$$v(\tau_{t,x}^{(n)},\ \xi_{t,x}(\tau_{t,x}^{(n)})) - v(t,\ \xi_{t,x}(t)) =$$
$$= \sum_{k=1}^{m} \int_t^{\tau_{t,x}^{(n)}} \sum_{i=1}^{m} \frac{\partial v}{\partial x^i}(s,\ \xi_{t,x}(s))\, b_{ki}(s,\ \xi_{t,x}(s))\, dw_k(s),$$

car $(s,\ \xi_{t,x}(s)) \in D$ pour $s < \tau_{t,x}$. En prenant l'espérance mathématique, on obtient en vertu de la formule (14), § 1,

$$v(t,\ x) = \mathsf{E}v(\tau_{t,x}^{(n)},\ \xi_{t,x}(\tau_{t,x}^{(n)})).$$

En passant à la limite pour $n \to \infty$ et en se servant de l'égalité

$$v(\tau_{t,x},\ \xi_{t,x}(\tau_{t,x})) = \psi(\tau_{t,x},\ \xi_{t,x}(\tau_{t,x}))$$

(le point $(\tau_{t,x},\ \xi_{t,x}(\tau_{t,x})) \in \Gamma$), on obtient ce qu'on voulait. $\blacksquare$

Considérons maintenant un processus homogène $\xi(t)$, solution de l'équation

$$d\xi(t) = a(\xi(t))\, dt + \sum_{k=1}^{m} b_k(\xi(t))\, dw_k(t) \qquad (2)$$

et défini pour $t \in [0, \infty[$. Soient $\xi_x(t)$ solution de l'équation (2) vérifiant la condition initiale $\xi(0) = x$, G domaine de $\mathcal{R}^m$, $D = [0, \infty[\times \mathcal{R}^m$. Dans ce cas, au lieu d'utiliser les solutions de l'équation $Lv = 0$ (de type parabolique) pour déterminer les quantités envisagées dans les théorèmes 1 et 2, on se servira des solutions d'une équation plus simple (de type elliptique) d'opérateur différentiel

$$L_1 u(x) = \sum_{i=1}^{m} a_i(x) \frac{\partial u(x)}{\partial x^i} + \frac{1}{2} \sum_{i,j=1}^{m} \frac{\partial^2 u(x)}{\partial x^i\, \partial x^j} \sum_{k=1}^{m} b_{ki}(x)\, b_{kj}(x). \qquad (3)$$

THÉORÈME 3. *Soient f fonction continue bornée dans l'adhérence de G, $u_\lambda(x)$ fonction définie, continue et bornée dans l'adhérence de G, deux fois continûment dérivable, nulle sur la frontière Γ de G et telle que*

$$\lambda u_\lambda(x) - L_1 u_\lambda(x) = f(x), \quad \lambda > 0. \tag{4}$$

Alors

$$u_\lambda(x) = \mathsf{E} \int_0^\infty e^{-\lambda t} f(\xi_x(t))\, \chi_{\{\tau_x > t\}}\, dt, \tag{5}$$

où τ_x est l'instant de première sortie du processus $\xi_x(t)$ du domaine G ($\tau_x = +\infty$ si $\xi_x(t) \in G$, $\forall t$).

Démonstration. Soit $t < \tau_x$. En appliquant la formule de Ito à la fonction $e^{-\lambda t} u_\lambda(\xi_x(t))$, on trouve

$$u_\lambda(\xi_x(t))\, e^{-\lambda t} - u_\lambda(x) = \int_0^t [-\lambda e^{-\lambda s} u_\lambda(\xi_x(s))\, ds +$$
$$+ e^{-\lambda s} L_1 u_\lambda(\xi_x(s))]\, ds +$$
$$+ \sum_{k=1}^m \int_0^t \sum_{i=1}^m \frac{\partial u_\lambda}{\partial x^i}(\xi_x(s))\, b_{ki}(\xi_x(s))\, dw_k(s). \tag{6}$$

Soit $\tau_x^{(n)}$ l'instant de première sortie du compact F_n, où F_n est une suite croissante de compacts telle que $\bigcup_n F_n = G$. Il vient $\tau_x^{(n)} \wedge T < \tau_x$ et les quantités $\frac{\partial u_\lambda}{\partial x^i}(\xi_x(s))$ sont bornées pour $s < \tau_x^{(n)} \wedge T$. En portant $t = \tau_x^{(n)} \wedge T$ dans (6) et eu égard à (4) on trouve

$$\mathsf{E} u_\lambda(\xi_x(\tau_x^{(n)} \wedge T))\, e^{-\lambda \tau_x^{(n)} \wedge T} - u_\lambda(x) = -\mathsf{E} \int_0^{\tau_x^{(n)} \wedge T} e^{-\lambda s} f(\xi_x(s))\, ds.$$

Un passage à la limite pour $n \to \infty$ donne

$$\mathsf{E} u_\lambda(\xi_x(\tau_x \wedge T))\, e^{-\lambda \tau_x \wedge T} - u_\lambda(x) = -\mathsf{E} \int_0^{\tau_x \wedge T} f(\xi_x(s))\, e^{-\lambda s}\, ds. \tag{7}$$

On remarquera que

$$\lim_{T \to \infty} u_\lambda(\xi(\tau_x \wedge T))\, e^{-\lambda \tau_x \wedge T} = 0,$$

car $u_\lambda(\xi_x(\tau_x))=0$ pour τ_x fini, puisque $\xi_x(\tau_x)\in\Gamma$ et

$$|u_\lambda(\xi_x(\tau_x\wedge T))|\,e^{-\lambda\tau_x\wedge T}\leqslant\sup_y|u_\lambda(y)|\,e^{-\lambda T}\to 0$$

pour $\tau_x=+\infty$.

Par ailleurs

$$|u_\lambda(\xi_x(\tau_x\wedge T))|\,e^{-\lambda\tau_x\wedge T}\leqslant\sup_y|u_\lambda(y)|,$$

donc

$$\lim_{T\to\infty}\mathsf{E}u_\lambda(\xi_x(\tau_x\wedge T))\,e^{-\lambda\tau_x\wedge T}=0.$$

Un passage à la limite pour $T\to\infty$ dans (7) achève la démonstration du théorème. ■

Pour trouver la répartition conjointe de τ et $\xi(\tau)$ il suffit de connaître $\mathsf{E}e^{-\lambda\tau}\psi(\xi(\tau))$ pour $\lambda>0$ et pour toutes les fonctions ψ suffisamment de fois dérivables, définies sur Γ.

THÉORÈME 4. *Soit $v_\lambda(x)$ fonction définie, bornée et continue sur l'adhérence de G, deux fois continûment dérivable dans G et solution de l'équation $\lambda v_\lambda(x)-L_1v_\lambda(x)=0$ $(\lambda\geqslant 0)$ dans G. Si $v_\lambda(x)=\psi(x)$ pour $x\in\Gamma$, alors*

$$v_\lambda(x)=\mathsf{E}e^{-\lambda\tau_x}\psi(\xi_x(\tau_x)).$$

Démonstration. En appliquant la formule de Ito à la fonction $e^{-\lambda t}v_\lambda(\xi_x(t))$, on trouve pour $t<\tau_x$

$$v_\lambda(\xi_x(t))\,e^{-\lambda t}-v_\lambda(x)=\sum_{k=1}^m\int_0^t\sum_{i=1}^m\frac{\partial v_\lambda(\xi_x(t))}{\partial x^i}\,b_{ki}(\xi_x(s))\,dw_k(s).$$

De là, exactement comme dans la démonstration du théorème 3, on déduit l'équation

$$\mathsf{E}v_\lambda(\xi_x(\tau_x))\,e^{-\lambda\tau_x}=v_\lambda(x).$$

Reste à noter que $v_\lambda(\xi_x(\tau_x))=\psi(\xi_x(\tau_x))$, puisque $\xi_x(\tau_x)\in\Gamma$. ■

REMARQUE 1. En posant $\lambda=0$ on peut trouver la répartition de $\xi(\tau)$: $\mathsf{E}\psi(\xi_x(\tau_x))=v(x)$, où $v(x)$ est une fonction continue bornée dans $G\cup\Gamma$, solution de l'équation $L_1v(x)=0$ dans G et telle que $v(x)=\psi(x)$ pour $x\in\Gamma$.

REMARQUE 2. Posons $\psi(x)=1$. On a alors $v_\lambda(x)=\mathsf{E}e^{-\lambda\tau_x}$. La fonction $v_\lambda(x)$ est continue, bornée dans $G\cup\Gamma$, vérifie la condition $v_\lambda(x)=1$ pour $x\in\Gamma$ et l'équation

$$\lambda v_\lambda(x)-L_1v_\lambda(x)=0. \tag{8}$$

Dérivons cette expression par rapport à λ et posons $\lambda = 0$. Soit $-\frac{\partial}{\partial\lambda}v_\lambda(x)|_{\lambda=0} = M_1(x)$. Il vient $M_1(x) = \mathsf{E}\tau_x$; de (8) et compte tenu de l'égalité $v_0(x) = 1$, il résulte

$$L_1M_1(x) = -1,$$

$M_1(x) = 0$ pour $x \in \Gamma$.

Si $M_1(x)$ est une fonction continue et bornée dans $G \cup \Gamma$, deux fois continûment dérivable dans G, nulle sur Γ et telle que $L_1M_1(x) = -1$, alors

$$\mathsf{E}\tau_x = M_1(x). \tag{9}$$

Pour le vérifier il faut appliquer la formule de Ito à la fonction $M_1(\xi_x(t))$ pour $t \in [0, \tau_x^{(n)}]$, soit

$M_1(\xi_x(\tau_x^{(n)})) - M_1(x) =$

$$= \int_0^{\tau_x^{(n)}} L_1M_1(\xi_x(s))\,ds + \sum_{k=1}^m \int_0^{\tau_x^{(n)}} \sum_{i=1}^m \frac{\partial}{\partial x^i} M_1(\xi_x(s))\, b_{ki}(\xi_x(s))\, dw_k(s) =$$

$$= -\tau_x^{(n)} + \sum_{k=1}^m \int_0^{\tau_x^{(n)}} \frac{\partial}{\partial x^i} M_1(\xi_x(s))\, b_{ki}(\xi_x(s))\, dw_k(s),$$

ensuite prendre l'espérance mathématique et passer à la limite pour $n \to \infty$.

Si la fonction $M_{n-1}(x) = \mathsf{E}(\tau_x)^{n-1}$ est continue dans $G \cup \Gamma$, et existe une fonction $M_n(x)$ continue dans $G \cup \Gamma$, deux fois dérivable dans G, nulle sur Γ et telle que

$$L_1M_n(x) = -nM_{n-1}(x),$$

alors

$$M_n(x) = \mathsf{E}(\tau_x)^n.$$

En effet, en procédant comme dans le cas précédent on obtient

$$M_n(x) = \mathsf{E}\int_0^{\tau_x} nM_{n-1}(\xi_x(s))\,ds.$$

On remarquera maintenant que $\tau_{\xi_x(s)} = \tau_x - s$ pour $s < \tau_x$ (à l'expiration du temps s le processus $\xi_x(\cdot)$ tombe dans le point $\xi_x(s)$ et le temps qu'il restera dans G sera de s inférieur à celui qu'il y passerait s'il y tombait à l'instant initial). Le processus étant markovien, on a donc pour $s < \tau_x$

$$\mathsf{E}[(\tau_x - s)^{n-1} | \mathfrak{F}_s] = M_{n-1}(\xi_x(s)),$$

où $\mathfrak{F}_s$ est la tribu engendrée par $w_k(u)$, $u \leqslant s$, $k = 1, \ldots, m$. Par suite,

$$M_n(x) = \mathsf{E} \int_0^{\tau_x} n\mathsf{E}\,[(\tau_x - s)^{n-1} \mid \mathfrak{F}_s]\, ds =$$

$$= n\mathsf{E} \int_0^{\infty} \chi_{\{\tau_x > s\}} \mathsf{E}\,[(\tau_x - s)^{n-1} \mid \mathfrak{F}_s]\, ds =$$

$$= n \int_0^{\infty} \mathsf{E}\chi_{\{\tau_x > s\}} \mathsf{E}\,[(\tau_x - s)^{n-1} \mid \mathfrak{F}_s]\, ds =$$

$$= n \int_0^{\infty} \mathsf{E}\chi_{\{\tau_x > s\}} (\tau_x - s)^{n-1}\, ds = \mathsf{E} \int_0^{\tau_x} n\,(\tau_x - s)^{n-1}\, ds = \mathsf{E}\,(\tau_x)^n.$$

Soit $\xi_x(t)$ processus homogène à une dimension, solution de l'équation stochastique

$$\xi_x(t) = x + \int_0^t a(\xi_x(s))\, ds + \int_0^t b(\xi_x(s))\, dw(s),$$

dont les coefficients vérifient les hypothèses du théorème 1, § 2. Soit τ_x l'instant de première sortie du processus $\xi_x(t)$ de l'intervalle $]\alpha, \beta[$. Considérons la fonction

$$\varphi(x) = \int_\alpha^x \exp\left\{-\int_\alpha^y \frac{2a(z)}{b^2(z)}\, dz\right\} dy.$$

Il est aisé de voir qu'elle est solution de l'équation

$$a(x)\,\varphi'(x) + \frac{1}{2} b^2(x)\,\varphi''(x) = L_1\varphi = 0$$

(car l'opérateur L_1 est précisément de cette forme). Toute solution de l'équation $Lu = 0$ s'écrit $u = c_1\varphi + c_2$, où c_1 et c_2 sont des constantes. La remarque 1 nous permet d'écrire

$$\mathsf{E}\varphi(\xi_x(\tau_x)) = \varphi(x).$$

Montrons que τ_x est fini. Cherchons pour cela la solution de l'équation

$$a(x)\, M_1'(x) + \frac{1}{2} b^2(x)\, M_1''(x) = -1, \quad M_1(\alpha) = M_1(\beta) = 0.$$

La solution de cette équation qui vérifie la condition $M_1(\alpha) = 0$ s'écrit

$$C\varphi(x) + 2\int_\alpha^x \frac{\varphi(x)-\varphi(z)}{b^2(z)\varphi'(z)}\,dz.$$

Comme $\varphi(\beta) \neq 0$, on trouve donc

$$M_1(x) = -2\frac{\varphi(x)}{\varphi(\beta)}\int_\alpha^\beta \frac{\varphi(\beta)-\varphi(z)}{b^2(z)\varphi'(z)}\,dz + 2\int_\alpha^x \frac{\varphi(x)-\varphi(z)}{b^2(z)\varphi'(z)}\,dz.$$

La remarque 2 indique que $\mathsf{E}\tau_x = M_1(x) < \infty$. La finitude de τ_x entraîne que $\xi(\tau_x)$ prend presque sûrement la valeur α ou β. Par suite

$$\varphi(x) = \mathsf{P}\{\xi(\tau_x) = \beta\}\,\varphi(\beta),$$

$$\mathsf{P}\{\xi(\tau_x) = \beta\} = \frac{\varphi(x)}{\varphi(\beta)},\quad \mathsf{P}\{\xi(\tau_x) = \alpha\} = \frac{\varphi(\beta)-\varphi(z)}{\varphi(\beta)}.$$

Soient $a = 0$, $b = 1$. Il vient $\varphi(x) = x - \alpha$,

$$\mathsf{P}\{\xi(\tau_x) = \alpha\} = \frac{\beta - x}{\beta - \alpha},\quad \mathsf{P}\{\xi(\tau_x) = \beta\} = \frac{x-\alpha}{\beta-\alpha}.$$

Dans ce cas il est facile de calculer tous les moments de τ_x. L'équation pour $M_1(x)$ s'écrit

$$\frac{1}{2}M_1''(x) = -1,\quad M_1(\alpha) = M_1(\beta) = 0,\quad M_1(x) = (x-\alpha)(x-\beta).$$

Pour $M_n(x)$ on a

$$\frac{1}{2}M_n''(x) = -nM_{n-1}(x),\quad M_n(\alpha) = M_n(\beta) = 0.$$

Donc

$$M_n(x) = 2n\int_\alpha^x (x-z)\,M_{n-1}(z)\,dz + 2n\frac{x-\alpha}{\beta-\alpha}\int_\alpha^\beta (\beta-z)\,M_{n-1}(z)\,dz.$$

Cette relation permet de calculer par récurrence tous les moments de τ_x. Quant à la répartition de τ_x on peut la trouver grâce à la remarque 2:

$$\mathsf{E}e^{-\lambda\tau_x} = v_\lambda(x),$$

où $v_\lambda(x)$ est la solution de l'équation

$$\frac{1}{2}v_\lambda''(x) - \lambda v_\lambda(x) = 0$$

qui vérifie les conditions aux limites $v_\lambda(\alpha) = v_\lambda(\beta) = 1$. Donc

$$v_\lambda(x) = \frac{\operatorname{ch}\left[\sqrt{2\lambda}\left(x - \frac{\alpha+\beta}{2}\right)\right]}{\operatorname{ch}\left[\sqrt{2\lambda}\,\frac{\beta-\alpha}{2}\right]}.$$

§ 6. Continuité absolue des mesures associées aux processus de diffusion

Si $\xi(t)$ est un processus aléatoire, défini sur $[0, T]$ et à valeurs dans $\mathcal{R}^m$, le théorème de Kolmogorov (chapitre II, § 2) lui associe une mesure sur l'espace $\{\mathcal{F}_{[0,T]}, \mathfrak{F}\}$, où $\mathcal{F}_{[0,T]}$ est l'espace des fonctions $x(t)$ définies sur $[0, T]$, à valeurs dans $\mathcal{R}^m$, $\mathfrak{F}$ une tribu engendrée par des ensembles cylindriques. Si l'on sait *a priori* que le processus $\xi(t)$ est continu, on peut construire la mesure qui lui est associée sur $\{\mathcal{C}_{[0,T]}, \mathfrak{F}\}$, où $\mathcal{C}_{[0,T]}$ est l'espace des fonctions continues. Dans la suite nous envisagerons exclusivement des processus continus et les mesures respectives seront définies sur $\mathcal{C}_{[0,T]}$. Soit μ_ξ mesure associée à un processus $\xi(\cdot)$. Nous allons nous intéresser aux conditions sous lesquelles les mesures associées à deux processus de diffusion $\xi_1(\cdot)$ et $\xi_2(\cdot)$ seront absolument continues l'une par rapport à l'autre.

On rappelle que si deux mesures μ_1 et μ_2 sont définies sur un espace probabilisable $\{X, \mathfrak{B}\}$, alors μ_2 est absolument continue par rapport à μ_1 si $\mu_2(B) = 0$ pour tous les $B \in \mathfrak{B}$ tels que $\mu_1(B) = 0$. Un théorème classique de Radon-Nikodym affirme que μ_2 est absolument continue par rapport à μ_1 si et seulement si existe une fonction $\rho(x)$ non négative $\mathfrak{B}$-mesurable telle que

$$\mu_2(B) = \int_B \rho(x)\,\mu_1(dx), \quad \forall B \in \mathfrak{B}.$$

La fonction $\rho(x)$ est définie de façon unique aux ensembles de mesure μ_1 nulle près; on l'appelle *densité (dérivée) de la mesure* μ_2 *par rapport à* μ_1 et on la note

$$\rho(x) = \frac{d\mu_2}{d\mu_1}(x).$$

Dans ce paragraphe on se propose d'étudier les conditions de continuité absolue des mesures associées aux processus de diffusion et de calculer aussi les densités respectives.

A signaler que si un processus $\xi(t) = \xi(t, \omega)$ est défini sur un espace probabilisé $\{\Omega, \mathfrak{S}, \mathsf{P}\}$, la mesure qui lui est associée est tout simplement l'image de P par l'application

$$\Omega \xrightarrow{\xi(\cdot,\omega)} \mathcal{C}_{[0,T]}.$$

Pour tout ensemble $B \in \mathfrak{F}$

$$\mu_\xi(B) = \mathsf{P}\{\omega : \xi(\cdot, \omega) \in B\}.$$

Ceci permet de construire sur $\mathcal{C}_{[0,T]}$ des mesures absolument continues par rapport à μ_ξ en les associant par l'application $\xi(\cdot,\omega)$

à la mesure

$$\widetilde{\mathsf{P}}(A) = \int_A \rho(\omega)\,\mathsf{P}(d\omega),$$

où $\rho(\omega)$ est une fonction non négative $\mathfrak{S}$-mesurable telle que

$$\int \rho(\omega)\,\mathsf{P}(d\omega) = 1$$

(cette dernière condition est nécessaire à la réalisation de $\widetilde{\mathsf{P}}(\Omega) = 1$). Soient $\tilde{\xi}(t)$ processus aléatoire défini par la fonction $\xi(t, \omega)$ sur l'espace probabilisé $\{\Omega, \mathfrak{S}, \widetilde{\mathsf{P}}\}$, $\mu_{\tilde{\xi}}$ la mesure associée à ce processus sur $\mathscr{C}_{[0,T]}$. Si $\mathfrak{S}^{\xi}$ est une tribu de Ω engendrée par $\xi(t, \omega)$, $t \in [0, T]$, alors

$$\frac{d\mu_{\tilde{\xi}}}{d\mu_{\xi}}(\xi(\cdot, \omega)) = \mathsf{E}(\rho(\omega) \mid \mathfrak{S}^{\xi}). \tag{1}$$

En effet, pour $A \in \mathfrak{F}$

$$\mu_{\tilde{\xi}}(A) = \widetilde{\mathsf{P}}(\xi(\cdot, \omega) \in A) = \int_A \rho(\omega)\,\mathsf{P}(d\omega) = \mathsf{E}\rho(\omega)\,\chi_A(\xi(\cdot, \omega)) =$$

$$= \mathsf{E}\chi_A(\xi(\cdot, \omega))\,\mathsf{E}(\rho(\omega) \mid \mathfrak{S}^{\xi}) = \int_A \mathsf{E}(\rho(\omega) \mid \mathfrak{S}^{\xi})\,|_{\xi = x(\cdot)}\mu_{\xi}(dx)$$

(on utilise le fait que $\mathsf{E}(\rho(\omega) \mid \mathfrak{S}^{\xi})$ est une fonction de $\xi(\cdot)$ et l'espérance mathématique de cette fonction, une intégrale par rapport à la mesure μ_{ξ}).

Nous allons prouver un théorème qui nous permettra de décrire une classe de fonctions $\rho(\omega)$ pour lesquelles la mesure image de $\widetilde{\mathsf{P}}$ sur $\mathscr{C}_{[0,T]}$ sera associée à un processus de diffusion, si seulement le processus $\xi(t, \omega)$ est processus de diffusion sur $\{\Omega, \mathfrak{S}, \mathsf{P}\}$.

THÉORÈME 1 (de Guirsanov). *Soient* $w_1(t), \ldots, w_m(t)$ *processus wienériens indépendants l'un de l'autre,* $\mathfrak{F}_t$, $0 \leqslant t \leqslant T$, *famille croissante de tribus;* $w_k(t)$ $\mathfrak{F}_t$-*mesurables, et* $\{w_k(s + t) - w_k(t),\ s > 0,\ k = 1, \ldots, m\}$ *indépendants dans leur ensemble de* $\mathfrak{F}_t$. *Si* $f_1(t), \ldots, f_m(t)$ *sont des fonctions* $\mathfrak{F}_t$-*mesurables sur* $[0, T]$, *et*

1) $$\int_0^T \sum_{k=1}^m f_k^2(t)\,dt < \infty,$$

2) $$\rho(\omega) = \mathsf{E}\exp\left\{\sum_{k=1}^m \int_0^T f_k(t)\,dw_k(t) - \frac{1}{2}\sum_{k=1}^m \int_0^T f_k^2(t)\,dt\right\},$$

3) $$\mathsf{E}\rho(\omega) = 1,$$

4) $$\widetilde{\mathsf{P}}(A) = \int_A \rho(\omega)\,\mathsf{P}(d\omega) = \mathsf{E}\chi_A(\omega)\,\rho(\omega),$$

alors les processus

$$\tilde{w}_k(t) = w_k(t) - \int_0^t f_k(s)\, dw_k(s)$$

sont des processus wienériens indépendants sur l'espace probabilisé $\{\Omega, \mathfrak{S}, \tilde{\mathsf{P}}\}$.

Démonstration. En cinq étapes : I à V.

I. LEMME 1. *Si* $\sum_{k=1}^{m} f_k^2(t) \leqslant N$, *alors*

$$\mathsf{E}\left(\exp\left\{\sum_{k=1}^{m}\int_t^T f_k(s)\, dw_k(s)\right\}\Big|\,\mathfrak{F}_t\right) \leqslant \exp\left\{\frac{1}{2} N(T-t)\right\}. \qquad (2)$$

Démonstration. Soient $f_k(s)$ fonctions en escalier sur $[t, T]$, $f_k(s) = f_k(s_j)$ pour $s \in [s_j, s_{j+1}]$, $t = s_0 < s_1 < \ldots < s_n = T$. L'égalité

$$\mathsf{E}\left(\exp\left\{\sum_{k=1}^{m} f_k(s_j)[w_k(s_{j+1}) - w_k(s_j)]\right\}\Big|\,\mathfrak{F}_{s_j}\right) =$$
$$= \exp\left\{\frac{1}{2}\sum_{k=1}^{m} f_k^2(s_j)(s_{j+1}-s_j)\right\} \leqslant \exp\left\{\frac{1}{2} N(s_{j+1}-s_j)\right\}$$

montre que

$$\mathsf{E}\left(\exp\left\{\sum_{j=0}^{n-1}\sum_{k=1}^{m} f_k(s_j)[w_k(s_{j+1}) - w_k(s_j)]\right\}\Big|\,\mathfrak{F}_t\right) \leqslant$$
$$\leqslant \exp\left\{\frac{1}{2} N(s_n - s_{n-1})\right\} \mathsf{E}\left(\exp\left\{\sum_{j=0}^{n-2}\sum_{k=1}^{m} f_k(s_j)[w_k(s_{j+1}) -\right.\right.$$
$$\left.\left. - w_k(s_j)]\right\}\Big|\,\mathfrak{F}_t\right) \leqslant \exp\left\{\frac{1}{2} N \sum_{j=0}^{n-1}(s_{j+1}-s_j)\right\} = \exp\left\{\frac{1}{2} N(T-t)\right\}.$$

Dans le cas général, (2) peut être déduite par passage à la limite.

II. LEMME 2. *Dans les hypothèses du lemme* 1,

$$\mathsf{E}\left(\exp\left\{\sum_{k=1}^{m}\int_t^T f_k(s)\, dw_k(s) - \frac{1}{2}\int_t^T \sum_{k=1}^{m} f_k^2(s)\, ds\right\}\Big|\,\mathfrak{F}_t\right) = 1.$$

Démonstration. Soit

$$\rho_{t_1, t_2}(\omega) = \exp\left\{\sum_{k=1}^{m}\int_{t_1}^{t_2} f_k(s)\, dw_k(s) - \frac{1}{2}\int_{t_1}^{t_2}\sum_{k=1}^{m} f_k^2(s)\, ds\right\}.$$

Pour $t_1=t_2$ on a $\rho_{t_1, t_2}(\omega)=1$. La formule de Ito donne

$$d\rho_{t_1, t}(\omega)=\rho_{t_1, t}(\omega) \sum_{k=1}^{m} f_k(t)\, dw_k(t).$$

Donc

$$\rho_{t_1, t_2}(\omega)=1+\int_{t_1}^{t_2} \rho_{t_1, s}(\omega) \sum_{k=1}^{m} f_k(s)\, dw_k(s).$$

Comme

$$\int_{t_1}^{t_2} \mathsf{E}\rho_{t_1,s}^2(\omega) \sum_{k=1}^{m} f_k^2(s)\, ds \leqslant N \int_{t_1}^{t_2} e^{2N(t_2-t_1)}\, ds<\infty,$$

en vertu du lemme, il vient

$$\mathsf{E}\left(\int_{t_1}^{t_2} \rho_{t_1, s}(\omega) \sum_{k=1}^{m} f_k(s)\, dw_k(s) \mid \mathfrak{F}_{t_1}\right)=0. \ \blacksquare$$

III. Corollaire 1. *Quelles que soient* $f_k(t)$, *fonctions* $\mathfrak{F}_t$*-mesurables telles que* $\int_0^T \sum_{k=1}^{m} f_k^2(t)\, dt<\infty$, *on a*

$$\mathsf{E}\left(\exp\left\{\sum_{k=1}^{m} \int_t^T f_k(s)\, dw_k(s)-\frac{1}{2} \int_t^T \sum_{k=1}^{m} f_k^2(s)\, ds\right\} \Big| \mathfrak{F}_t\right) \leqslant 1.$$

Cette inégalité est une conséquence du théorème de Fatou.

IV. Corollaire 2. *Si* $\mathsf{E}\rho_{0, T}(\omega)=1$, *alors*

$$\mathsf{E}(\rho_{t_1, t_2}(\omega) \mid \mathfrak{F}_{t_1})=1 \quad \textit{pour} \quad 0 \leqslant t_1<t_2 \leqslant T.$$

En effet,

$$1=\mathsf{E}\rho_{0, t_1}(\omega)\, \mathsf{E}(\rho_{t_1, t_2}(\omega)\, \mathsf{E}(\rho_{t_2, T}(\omega) \mid \mathfrak{F}_{t_2}) \mid \mathfrak{F}_{t_1}). \tag{3}$$

Si l'inégalité $\mathsf{E}(\rho_{t_1,t_2}(\omega) \mid \mathfrak{F}_{t_1})<1$ était réalisée avec une probabilité positive, alors en vertu du corollaire 1 le second membre (3) serait strictement inférieur à 1.

V. On désignera par $\tilde{\mathsf{E}}$ l'espérance mathématique par rapport à la mesure $\tilde{\mathsf{P}}$. Pour toute variable aléatoire $\xi(\omega)$ définie sur $\{\Omega, \mathfrak{S}\}$, on a

$$\tilde{\mathsf{E}}\xi(\omega)=\mathsf{E}\xi(\omega)\, \rho(\omega).$$

Pour prouver le théorème il suffit de montrer que

$$\tilde{\mathsf{E}}\left(\exp\left\{i\sum_{k=1}^{m}(\tilde{w}_k(t_2)-\tilde{w}_k(t_1))z_k\right\}\Big|\,\mathfrak{F}_{t_1}\right)=$$

$$=\exp\left\{-\frac{1}{2}(t_2-t_1)\sum_{k=1}^{m}z_k^2\right\}\qquad(4)$$

pour $t_1<t_2$ et $\forall z_k$ réels. Autrement dit, pour toute variable η $\mathfrak{F}_{t_1}$-mesurable bornée on a

$$\tilde{\mathsf{E}}\eta\exp\left\{i\sum_{k=1}^{m}(\tilde{w}_k(t_2)-\tilde{w}_k(t_1))z_k\right\}=\exp\left\{-\frac{1}{2}(t_2-t_1)\sum_{k=1}^{m}z_k^2\right\}\tilde{\mathsf{E}}\eta.$$

(4) bis

Pour toute variable ξ $\mathfrak{F}_t$-mesurable il résulte du corollaire 2

$$\tilde{\mathsf{E}}\xi=\mathsf{E}\xi\rho_{0,t}(\omega)\rho_{t,T}(\omega)=\mathsf{E}\xi\rho_{0,t}(\omega)\,\mathsf{E}(\rho_{t,T}(\omega)\,|\,\mathfrak{F}_t)=\mathsf{E}\xi\rho_{0,t}(\omega),$$

donc (4) est équivalent à

$$\mathsf{E}\eta'\exp\left\{i\sum_{k=1}^{m}(\tilde{w}_k(t_2)-\tilde{w}_k(t_1))z_k\right\}\rho_{t_1,t_2}(\omega)=$$

$$=\exp\left\{-\frac{1}{2}(t_2-t_1)\sum_{k=1}^{m}z_k^2\right\}\mathsf{E}\eta',\qquad(5)$$

où $\eta'=\eta\rho_{0,t_1}(\omega)$ est une variable $\mathfrak{F}_{t_1}$-mesurable telle que $\mathsf{E}|\eta'|<\infty$.

La formule de Ito donne

$$d\exp\left\{i\sum_{k=1}^{m}(\tilde{w}_k(t)-\tilde{w}_k(t_1))z_k\right\}\rho_{t_1,t}(\omega)=$$

$$=\exp\left\{i\sum_{k=1}^{m}(\tilde{w}_k(t)-\tilde{w}_k(t_1))z_k\right\}\rho_{t_1,t}(\omega)\times$$

$$\times\left[\sum_{k=1}^{m}f_k(t)\,dw_k(t)+i\sum_{k=1}^{m}z_k\,dw_k(t)-\frac{1}{2}\sum_{k=1}^{m}z_k^2\,dt+i\sum_{k=1}^{m}f_k(t)\,z_k\,dt\right].$$

Par suite,

$$\exp\left\{i\sum_{k=1}^{m}(\tilde{w}_k(t_2)-\tilde{w}_k(t_1))z_k\right\}\rho_{t_1,t_2}(\omega)=$$

$$=1+\int_{t_1}^{t_2}\exp\left\{i\sum_{k=1}^{m}(\tilde{w}_k(t)-\tilde{w}_k(t_1))z_k\right\}\rho_{t_1,t}(\omega)\times$$

$$\times\left[i\sum_{k=1}^{m} z_k\,dw_k(t)+\sum_{k=1}^{m} f_k(t)\,dw_k(t)\right]-$$

$$-\frac{1}{2}\sum_{k=1}^{m} z_k^2\int_{t_1}^{t_2}\exp\left\{i\sum_{k=1}^{m}(\tilde{w}_k(t)-\tilde{w}_k(t_1))\,z_k\right\}\rho_{t_1,\,t}(\omega)\,dt.$$

Soit $\sum_{k=1}^{m} f_k^2(t)\leqslant N$. Le lemme 1 nous dit que

$$\mathsf{E}\left(\int_{t_1}^{t_2}\exp\left\{i\sum_{k=1}^{m}(\tilde{w}_k(t)-\tilde{w}_k(t_1))\,z_k\right\}\rho_{t_1,\,t_2}(\omega)\times\right.$$

$$\left.\times\sum_{k=1}^{m}(iz_k+f_k(t))\,dw_k(t)\,|\,\mathfrak{F}_{t_1}\right)=0.$$

Donc

$$\mathsf{E}\eta'\exp\left\{i\sum_{k=1}^{m}(\tilde{w}_k(t_2)-\tilde{w}_k(t_1))\,z_k\right\}\rho_{t_1,\,t_2}(\omega)=$$

$$=\mathsf{E}\eta'-\frac{1}{2}\sum_{k=1}^{m} z_k^2\int_{t_1}^{t_2}\mathsf{E}\eta'\exp\left\{i\sum_{k=1}^{m}(\tilde{w}_k(t)-\tilde{w}_k(t_1))\,z_k\right\}\rho_{t_1,\,t}(\omega)\,dt.$$

En posant

$$\mathsf{E}\eta'\exp\left\{i\sum_{k=1}^{m}(\tilde{w}_k(t_2)-\tilde{w}_k(t_1))\,z_k\right\}\rho_{t_1,\,t_2}(\omega)=\lambda(t_2),$$

on trouve

$$\lambda(t_2)=\mathsf{E}\eta'-\frac{1}{2}\sum_{k=1}^{m} z_k^2\int_{t_1}^{t_2}\lambda(t)\,dt,$$

d'où

$$\lambda(t_2)=\mathsf{E}\eta'\exp\left\{-\frac{1}{2}\sum_{k=1}^{m} z_k^2(t_2-t_1)\right\}.$$

Ceci démontre la formule (5) sous les hypothèses faites sur $f_k(t)$.

Supposons maintenant que $f_k^{(N)}(t)$ vérifient les conditions

$$\sum_{k=1}^{m}[f_k^{(N)}(t)]^2\leqslant N$$

et

$$\int_0^T \sum_{k=1}^m (f_k(t) - f_k^{(N)}(t))^2\,dt \to 0$$

pour $N \to \infty$. Posons

$$\widetilde{w}_k^{(N)}(t) = w_k(t) - \int_0^t f_k^{(N)}(s)\,ds,$$

$$\rho_{t_1, t_2}^{(N)}(\omega) = \exp\left\{ \sum_{k=1}^m \int_{t_1}^{t_2} f_k^{(N)}(t)\,dw_k(t) - \frac{1}{2}\int_{t_1}^{t_2} \sum_{k=1}^m (f_k^{(N)}(t))^2\,dt \right\}.$$

On a démontré que

$$\mathsf{E}\eta\rho_{0, t_2}^{(N)}(\omega) \exp\left\{ i \sum_{k=1}^m (\widetilde{w}_k^{(N)}(t_2) - \widetilde{w}_k^{(N)}(t_1))\,z_k \right\} =$$

$$= \mathsf{E}\eta\rho_{0, t_1}^{(N)}(\omega) \exp\left\{ -\frac{1}{2} \sum_{k=1}^m z_k^2 (t_2 - t_1) \right\}. \quad (6)$$

Comme $\widetilde{w}_k^{(N)}(t) \to \widetilde{w}_k(t)$, $k = 1, \ldots, m$, stochastiquement, on a

$$\lim_{N \to \infty} \mathsf{E}\eta\rho_{0, t_2}(\omega) \exp\left\{ i \sum_{k=1}^m (\widetilde{w}_k^{(N)}(t_2) - \widetilde{w}_k^{(N)}(t_1))\,z_k \right\} =$$

$$= \mathsf{E}\eta\rho_{0, t_2}(\omega) \exp\left\{ i \sum_{k=1}^m (\widetilde{w}_k(t_2) - \widetilde{w}_k(t_1))\,z_k \right\}. \quad (7)$$

Ensuite,

$$\left| \mathsf{E}\eta\rho_{0, t_2}^{(N)}(\omega) \exp\left\{ i \sum_{k=1}^m (\widetilde{w}_k^{(N)}(t_2) - \widetilde{w}_k^{(N)}(t_1))\,z_k \right\} - \right.$$

$$\left. - \mathsf{E}\eta\rho_{0, t_2}(\omega) \exp\left\{ i \sum_{k=1}^m (\widetilde{w}_k^{(N)}(t_2) - \widetilde{w}_k^{(N)}(t_1))\,z_k \right\} \right| \leqslant$$

$$\leqslant c\mathsf{E}\,|\rho_{0, t_2}^{(N)}(\omega) - \rho_{0, t_2}(\omega)|, \quad (8)$$

$$|\mathsf{E}\eta\rho_{0, t_1}^{(N)}(\omega) - \mathsf{E}\eta\rho_{0, t_1}(\omega)| \leqslant c\mathsf{E}\,|\rho_{0, t_1}^{(N)}(\omega) - \rho_{0, t_1}(\omega)|, \quad (9)$$

où c est tel que $|\eta| \leqslant c$. Prouvons que

$$\lim_{N \to \infty} \mathsf{E}\,|\rho_{0, t}^{(N)}(\omega) - \rho_{0, t}(\omega)| = 0. \quad (10)$$

On a

$$\mathsf{E}\,|\rho_{0, t}^{(N)}(\omega) - \rho_{0, t}(\omega)| = \mathsf{E}\,(|\rho_{0, t}^{(N)}(\omega) - \rho_{0, t}(\omega)| + \rho_{0, t}(\omega) - \rho_{0, t}^{(N)}(\omega)),$$

puisque

$$\mathsf{E}\rho_{0,t}(\omega) = \mathsf{E}\rho_{0,t}^{(N)}(\omega) = 1.$$

Or

$$|\rho_{0,t}^{(N)}(\omega) - \rho_{0,t}(\omega)| + \rho_{0,t}(\omega) - \rho_{0,t}^{(N)}(\omega) \leqslant 2\rho_{0,t}(\omega)$$

et

$$\rho_{0,t}^{(N)}(\omega) - \rho_{0,t}(\omega) \to 0$$

stochastiquement. Donc (10) est réalisée d'après le théorème de Lebesgue. En passant à la limite dans (6) et en tenant compte de (7), (8), (9) et (10), on obtient (5) (eu égard à η'). ■

Appliquons ce théorème à la démonstration de la continuité absolue des mesures associées aux deux processus de diffusion, solutions des équations différentielles stochastiques

$$d\xi_i(t) = a_i(t, \xi_i(t))\,dt + \sum_{k=1}^{m} (b_k(t, \xi_i(t))\,dw_k(t), \quad \xi_i(0) = x. \tag{11}$$

THÉORÈME 2. *Supposons que les coefficients des équations* (*11*) *vérifient les conditions*:

I. *Il existe K tel que*

1) $$|a_1(t, x) - a_1(t, y)| + |a_2(t, x) - a_2(t, y)| + \\ + \sum_{k=1}^{m} |b_k(t, x) - b_k(t, y)| \leqslant K|x - y|;$$

2) $$|a_1(t, x)|^2 + |a_2(t, x)|^2 + \sum_{k=1}^{m} |b_k(t, x)|^2 \leqslant K(1 + |x|^2).$$

II. *Il existe des fonctions continues $\lambda_1(t, x), \ldots, \lambda_m(t, x)$ telles que*

$$a_2(t, x) - a_1(t, x) = \sum_{k=1}^{m} \lambda_k(t, x)\, b_k(t, x).$$

Alors la mesure μ_{ξ_2} est absolument continue par rapport à μ_{ξ_1} et

$$\frac{d\mu_{\xi_2}}{d\mu_{\xi_1}}(\xi_1(\cdot, \omega)) = \\ = \mathsf{E}\left[\exp\left\{\sum_{k=1}^{m} \int_0^T \lambda_k(s, \xi_1(s))\,d\omega_k(s) - \frac{1}{2}\sum_{k=1}^{m} \int_0^T \lambda_k^2(s, \xi_1(s))\,ds\right\} \Big| \mathfrak{S}^{\xi_1}\right], \tag{12}$$

où $\mathfrak{S}^{\xi_1}$ est une tribu engendrée par $\xi_1(t, \omega)$, $t \in [0, T]$.

Démonstration. Soient $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ l'espace probabilisé sur lequel sont définis les processus $w_k(t)$, $k = 1, \ldots, m$, et $\xi_1(t)$

solution de l'équation (11) pour $i = 1$. Posons

$$\rho(\omega) = \exp\left\{\sum_{k=1}^{m} \int_0^T \lambda_k(s, \xi_1(s))\, dw_k(s) - \frac{1}{2} \sum_{k=1}^{m} \int_0^T \lambda_k^2(s, \xi_1(s))\, ds\right\}.$$

Supposons d'abord que

$$\mathsf{E}\rho(\omega) = 1 \tag{13}$$

(ceci aura lieu par exemple si $\lambda_k(s, x)$ sont bornées en vertu du lemme 2).

Soit $\widetilde{\mathsf{P}}$ mesure sur $\mathfrak{S}$ définie par

$$\widetilde{\mathsf{P}}(A) = \int_A \rho(\omega)\, \mathsf{P}(d\omega).$$

Le théorème 1 affirme que les processus

$$\widetilde{w}_k(t) = w_k(t) - \int_0^t \lambda_k(s, \xi_1(s))\, ds$$

sont des processus wienériens indépendants. Considérons le processus $\xi_1(t)$ sur l'espace probabilisé $\{\Omega, \mathfrak{S}, \mathsf{P}\}$. Il vérifie:

$$\xi_1(t) - x_0 = \int_0^t a_1(s, \xi_1(s))\, ds + \sum_{k=1}^{m} \int_0^t b_k(s, \xi_1(s))\, dw_k(s) =$$

$$= \int_0^t a_1(s, \xi_1(s))\, ds + \sum_{k=1}^{m} \int_0^t b_k(s, \xi_1(s))\, [d\widetilde{w}_k(s) + \lambda_k(s, \xi_1(s))\, ds] =$$

$$= \int_0^t \left[a_1(s, \xi_1(s)) + \sum_{k=1}^{m} \lambda_k(s, \xi_1(s))\, b_k(s, \xi_1(s))\right] ds +$$

$$+ \sum_{k=1}^{m} \int_0^t b_k(s, \xi_1(s))\, d\widetilde{w}_k(s) = \int_0^t a_2(s, \xi_1(s))\, ds +$$

$$+ \sum_{k=1}^{m} \int_0^t b_k(s, \xi_1(s))\, d\widetilde{w}_k(s).$$

Il est donc confondu avec la solution de l'équation

$$d\widetilde{\xi}_2(t) = a_2(t, \widetilde{\xi}_2(t))\, dt + \sum_{k=1}^{m} b_k(t, \widetilde{\xi}_2(t))\, d\widetilde{w}_k(t)$$

sur l'espace probabilisé $\{\Omega, \mathfrak{S}, \tilde{\mathsf{P}}\}$. Par suite la mesure $\mu_{\tilde{\xi}_2}$ est confondue avec μ_{ξ_2}. En vertu de la formule (1)

$$\frac{d\mu_{\xi_2}}{d\mu_{\xi_1}}(\xi_1(\cdot,\ \omega)) = \mathsf{E}(\rho(\omega) \mid \mathfrak{S}^{\xi_1}),$$

le théorème est prouvé sous l'hypothèse (13).

Supposons maintenant que (13) est toujours réalisée. Soient $\lambda_k^{(N)}(t,\ x)$ tels que $\lambda_k(t,\ x) = \lambda_k^{(N)}(t,\ x)$ pour $|x| \leqslant N$, $\lambda_k^{(N)}(t,\ x)$ continus, $\lambda_k^{(N)}(t,\ x) = 0$ pour $|x| > 2N$,

$$a_2^{(N)}(t,\ x) = a_1(t,\ x) + \sum_{k=1}^{m} \lambda_k^{(N)}(t,\ x)\, b_k(t,\ x)$$

une fonction telle que

$$|a_2^{(N)}(t,\ x) - a_2^{(N)}(t,\ y)| \leqslant K_1 |x-y|,$$
$$|a_2^{(N)}(t,\ x)|^2 \leqslant K_1(1+|x|^2),$$

où K_1 est une constante (ne dépendant pas de N). Alors, en posant

$$\rho_N(\omega) = \exp\left\{\sum_{k=1}^{m} \int_0^T \lambda_k^{(N)}(s,\ \xi_1(s))\, dw_k(s) - \right.$$
$$\left. - \frac{1}{2}\sum_{k=1}^{m} \int_0^T [\lambda_k^{(N)}(s,\ \xi_1(s))]^2\, ds\right\},$$

on aura, d'après ce qui a été démontré,

$$\frac{d\mu_{\xi_2^{(N)}}}{d\mu_{\xi_1}}(\xi_1(\cdot,\ \omega)) = \mathsf{E}(\rho_N(\omega) \mid \mathfrak{S}^{\xi_1}),$$

où $\xi_2^{(N)}(t)$ est solution de l'équation différentielle stochastique

$$\xi_2^{(N)}(t) = x_0 + \int_0^t a_2^{(N)}(s,\ \xi_2^{(N)}(s))\, ds + \sum_{k=1}^{m} \int_0^t b_k(s,\ \xi_2^{(N)}(s))\, dw_k(s). \quad (14)$$

Si l'on remarque maintenant que $\rho_N(\omega) = \rho(\omega)$ pour $\sup_t |\xi_1(t,\ \omega)| \leqslant N$, on aura

$$\mathsf{E}\rho(\omega)\,\chi\{\sup_t |\xi_1(t,\ \omega)| \leqslant N\} = \mathsf{E}\rho_N(\omega)\,\chi\{\sup_t |\xi_1(t,\ \omega)| \leqslant N\} =$$
$$= \mathsf{P}\{\sup_t |\xi_2^{(N)}(t)| \leqslant N\}. \quad (15)$$

Evaluons $\mathsf{E}\,|\,\xi_2^{(N)}(t)\,|^2$. Comme

$$\mathsf{E}\,|\,\xi_2^{(N)}(t)\,|^2 \leqslant (m+2)\left[\,|\,x_0\,|^2 + \mathsf{E}\left|\int_0^t a_2^{(N)}(s,\ \xi_2^{(N)}(s))\,ds\right|^2 + \right.$$

$$\left. + \mathsf{E}\sum_{k=1}^m \left|\int_0^t b_k(s,\ \xi_2^{(N)}(s))\,dw_k(s)\right|^2\right] =$$

$$= (m+2)\left[\,|\,x_0\,|^2 + \mathsf{E}\int_0^t \left[\,|\,a_2^{(N)}(s,\ \xi_2^{(N)}(s))\,|^2 + \right.\right.$$

$$\left.\left. + \sum_{k=1}^m |\,b_k(s,\ \xi_2^{(N)}(s))\,|^2\right] ds\right] \leqslant$$

$$\leqslant (m+2)\left[\,|\,x_0|^2 + (TK_1 + mK)\int_0^t (1 + \mathsf{E}\,|\,\xi_2^{(N)}(s)\,|^2)\,ds\right],$$

le lemme 1, § 2, affirme l'existence de K_2 (ne dépendant pas de N) tel que

$$\mathsf{E}\,|\,\xi_2^{(N)}(t)\,|^2 \leqslant K_2. \tag{16}$$

On a

$$\mathsf{P}\{\sup_t |\,\xi_2^{(N)}(t)\,| > N\} \leqslant \mathsf{P}\left\{\int_0^T |\,a_2^{(N)}(s,\ \xi_2^{(N)}(s))\,|\,ds + \right.$$

$$\left. + \sum_{k=1}^m \sup_{t\leqslant T}\left|\int_0^t b_k(s,\ \xi_2^{(N)}(s))\,dw_k(s)\right| > N - |\,x_0\,|\right\}.$$

L'inégalité de Tchébychev, la majoration

$$\int_0^T |\,a_2^{(N)}(s,\ \xi_2^{(N)}(s))\,|\,ds \leqslant \int_0^T \frac{1}{2}\,(1 + |\,a_2^{(N)}(s,\ \xi_2^{(N)}(s))\,|^2)\,ds \leqslant$$

$$\leqslant \frac{T}{2} + \frac{1}{2}K_1\int_0^T (1 + |\,\xi_2^{(N)}(s)\,|^2)\,ds,$$

la propriété V, § 1, en vertu de laquelle

$$\mathsf{P}\left\{\sup_{t\leqslant T}\left|\int_0^t b_k(s,\ \xi_2^{(N)}(s))\,dw_k(s)\right| > c\right\} \leqslant$$

$$\leqslant \frac{1}{c^2}\int_0^T \mathsf{E}\,|\,b_k(s,\ \xi_2^{(N)}(s))\,|^2\,ds \leqslant \frac{K}{c^2}\int_0^T \mathsf{E}\,(1 + |\,\xi_2^{(N)}(s)\,|^2)\,ds$$

et enfin (16) montrent que

$$\mathsf{P}\{\sup_t |\xi_2^{(N)}(t)| > N\} = O\left(\frac{1}{N^2}\right),$$

et par suite le second membre de (15) tend vers 1 pour $N \to \infty$. Donc

$$\mathsf{E}\rho(\omega) = \lim_{N\to\infty} \mathsf{E}\rho(\omega)\chi\{\sup_t |\xi_1(t, \omega)| \leqslant N\} = 1.$$

Ce qui démontre (13) et le théorème. ■

Remarque. Supposons que la matrice $B(s, x) = (b_{kj}(s, x))$, $j, k = 1, \ldots, m$ (b_{kj} sont les coordonnées du vecteur b_k) est non dégénérée. Alors dans les hypothèses du théorème 2

$$\frac{d\mu_{\xi_2}}{d\mu_{\xi_1}}(\xi_1(\cdot, \omega)) = \rho(\omega).$$

Pour le vérifier nous allons montrer que $\rho(\omega)$ est une variable $\mathfrak{S}^{\xi_1}$-mesurable. Il nous suffit de prouver que les processus $w_k(t)$, $k = 1, \ldots, m$, sont $\mathfrak{S}^{\xi_1}$-mesurables. Soient $c_{kj}(s, x)$ les éléments de l'inverse de la matrice $B(s, x)$:

$$\sum_{j=1}^{m} b_{kj}(s, x)\, c_{ji}(s, x) = \delta_{ki}.$$

Soient d'autre part $0 = s_0 < s_1 < \ldots < s_n = t$, $\Delta s_l = s_{l+1} - s_l$; il vient

$$w_k(t) = \lim_{\max \Delta s_l \to 0} \sum_{l=0}^{n-1} \sum_{j=1}^{m} c_{jk}(s_l, \xi_1(s_l)) \times$$
$$\times \left[\xi_1^j(s_{l+1}) - \xi_1^j(s_l) - \int_{s_l}^{s_{l+1}} a_1^j(s, \xi_1(s))\, ds\right] \quad (17)$$

au sens de la convergence stochastique; ici ξ_1^j et a_1^j sont les composantes des vecteurs ξ_1 et a_1. En effet, l'expression figurant sous le signe de la limite s'écrit

$$\sum_{l=0}^{n-1} \sum_{j=1}^{m} c_{jk}(s_l, \xi(s_l)) \sum_{i=1}^{m} \int_{s_l}^{s_{l+1}} b_{ij}(s, \xi_1(s))\, dw_i(s) =$$
$$= \sum_{l=0}^{n-1} \sum_{j=1}^{m} \sum_{i=1}^{m} \int_{s_l}^{s_{l+1}} c_{jk}(s_l, \xi_1(s_l))\, b_{ij}(s, \xi_1(s))\, dw_i(s) =$$
$$= \sum_{i=1}^{m} \int_0^t \psi_i^{(n)}(s)\, dw_i(s),$$

où

$$\psi_i^{(n)}(s) = \sum_{j=1}^{m} b_{ij}(s, \xi_1(s))\, c_{jk}(s_l, \xi_1(s_l)) \text{ pour } s_l \leqslant s < s_{l+1}.$$

On remarquera que $\psi_i^{(n)}(s) - \delta_{ik} \to 0$ uniformément avec max Δs_l en vertu de la continuité des fonctions $b_{ij}(s, \xi_1(s))$ et $c_{jk}(s, \xi_1(s))$. Donc (17) résulte de la propriété IV, § 1, pour les intégrales stochastiques. Le second membre de (17) est visiblement $\mathfrak{S}^{\xi_1}$-mesurable.

CHAPITRE IX

THÉORÈMES LIMITES POUR PROCESSUS ALÉATOIRES

A maintes reprises nous avons eu affaire à des processus déduits de processus plus simples par passage à la limite.

La théorie des processus aléatoires fait une large part aux méthodes de recherche de la répartition des fonctionnelles d'un processus aléatoire, par exemple:

$$\int_{t_1}^{t_2} f(\xi(s))\,ds, \quad \sup_{t_1 \leqslant t \leqslant t_2} \xi(t), \quad \inf_{t_1 \leqslant t \leqslant t_2} \xi(t).$$

Une question émerge tout naturellement: si l'on obtient un processus $\xi(t)$ par passage à la limite sur une suite de processus $\xi_n(t)$, ne pourrait-on pas obtenir aussi la répartition des fonctionnelles du processus $\xi(t)$ sachant celles des fonctionnelles des processus $\xi_n(t)$?

On admettra que la suite de processus $\xi_n(t)$ est au moins faiblement convergente vers un processus $\xi(t)$, c'est-à-dire les répartitions finidimensionnelles de $\xi_n(t)$ convergent vers les répartitions finidimensionnelles de $\xi(t)$. Ces hypothèses sont trop faibles pour qu'on puisse en déduire la convergence des répartitions pour une classe assez large de fonctionnelles (celles mentionnées plus haut par exemple). Il faudra donc chercher d'autres conditions sous lesquelles les répartitions des fonctionnelles (d'une classe F) de processus $\xi_n(t)$ convergeront vers les répartitions des fonctionnelles correspondantes du processus $\xi(t)$. La classe F doit être telle que $f(\xi_n(t))$ et $f(\xi(t))$ ($f \in F$) soient des variables aléatoires. Donc le choix de la classe F est tributaire des propriétés des processus $\xi_n(t)$ et $\xi(t)$.

Nous traiterons les cas où les processus sont presque sûrement continus ou présentent presque sûrement des discontinuités de seconde espèce.

La classe de fonctionnelles considérée ne sera pas la même dans les deux cas.

Les théorèmes limites pour processus aléatoires sont importants pas seulement pour la détermination des répartitions de fonctionnelles d'un processus limite moyennant un passage à la limite sur

des processus plus simples. Il est tout aussi naturel d'utiliser les processus continus pour décrire le comportement à la limite des processus discrets : les processus à accroissements indépendants pour la description d'une suite de sommes de variables aléatoires indépendantes, les processus markoviens continus pour la description d'une chaîne de Markov à temps discret. Dans ce cas on étudiera le passage à la limite de processus, variant à des dates fixes, vers des processus continus, sachant que les intervalles séparant ces dates tendent vers zéro. On étudiera également des critères de convergence faible de tels processus.

§ 1. Convergence faible de répartitions dans un espace métrique

Supposons que les réalisations des processus $\xi_n(t)$ et $\xi(t)$ $(t \in [a, b])$ appartiennent presque sûrement à un espace métrique fonctionnel X de métrique $\rho_X(x, y)$, $x, y \in X$. Si, par exemple, $\xi_n(t)$ et $\xi(t)$ sont presque sûrement continus, ils appartiendront à l'espace $\mathscr{C}$ des fonctions continues muni de la métrique

$$\rho_{\mathscr{C}}(x, y) = \sup_t |x(t) - y(t)|.$$

La classe F de fonctionnelles pour lesquelles on cherche les critères de convergence de la répartition de $f(\xi_n(t))$ vers la répartition de $f(\xi(t))$ sera une collection de fonctions continues sur X dans la métrique ρ_X. Pour que $f(\xi(t))$ soit variable aléatoire, il suffit d'exiger la séparabilité de l'ensemble X et la mesurabilité de l'ensemble $\{\omega: \xi(t) \in S\}$ par rapport à l'espace probabilisé initial pour toute boule ouverte S de X (car dans ce cas l'ensemble $\{\omega : \xi(t) \in A\}$ sera mesurable pour tout borélien A de X).

Dans la suite on admettra que X est un espace séparable et les processus $\xi_n(t)$ et $\xi(t)$ réunissent les conditions formulées plus haut. Au processus aléatoire $\xi(t)$ (resp. $\xi_n(t)$) on associe une mesure μ (resp. μ_n) définie sur la tribu des boréliens $\mathfrak{B}$ de X par :

$$\mu(A) = \mathsf{P}\{\xi(\cdot) \in A\} \text{ (resp. } \mu_n(A) = \mathsf{P}\{\xi_n(\cdot) \in A\}).$$

Toute fonctionnelle f $\mathfrak{B}$-mesurable vérifie la relation

$$\mathsf{E} f(\xi(t)) = \int f(x)\, \mu(dx).$$

A noter qu'une condition nécessaire et suffisante pour que les répartitions de $f(\xi_n(t))$ convergent vers la répartition de $f(\xi(t))$ pour toutes les fonctionnelles continues est que

$$\lim_{n\to\infty} \int f(x)\, \mu_n(dx) = \int f(x)\, \mu(dx), \tag{1}$$

pour toutes les fonctionnelles continues bornées f.

En effet, la convergence de la répartition de $f(\xi_n(t))$ vers celle de $f(\xi(t))$ et le fait que f est bornée entraînent la convergence de $\mathsf{E}f(\xi_n(t))$ vers $\mathsf{E}f(\xi(t))$ et partant la formule (1). Par ailleurs, de (1) il suit que la fonction caractéristique de $f(\xi_n(t))$, $\forall f$, converge vers celle de $f(\xi(t))$:

$$\lim_{n\to\infty} \mathsf{E}e^{i\lambda f(\xi_n(t))} = \lim_{n\to\infty} \int e^{i\lambda f(x)}\mu_n(dx) = \int e^{i\lambda f(x)}\mu(dx) = \mathsf{E}e^{i\lambda f(\xi(t))}.$$

D é f i n i t i o n. *Si la relation* (1) *est réalisée pour toutes les fonctions $f(x)$ continues bornées, on dit que la suite μ_n converge faiblement vers μ et l'on note $\mu_n \Rightarrow \mu$.*

D é f i n i t i o n. *Une suite de mesures μ_n est faiblement compacte si de l'une quelconque de ses sous-suites on peut extraire une suite de mesures faiblement convergente.*

THÉORÈME 1. *Soient X espace séparable complet, $\mathfrak{B}$ la tribu de ses boréliens. Pour qu'une suite de mesures μ_n sur $\mathfrak{B}$ soit faiblement compacte il est nécessaire et suffisant que*

a) $\sup_n \mu_n(X) < \infty$;

b) $\forall \varepsilon > 0$, $\exists K$ *compact tel que* $\sup_n \mu_n(X \setminus K) < \varepsilon$.

On aura besoin du lemme suivant.

LEMME 1. *Si X est un compact et $\sup_n \mu_n(X) = H < \infty$, alors la suite de mesures μ_n est faiblement compacte.*

D é m o n s t r a t i o n. Soit $\mathscr{C}_X$ l'espace des fonctions continues f sur X, $\|f\| = \sup_{x\in X} |f(x)|$. L'espace $\mathscr{C}_X$ est un espace polonais normé. Soit $\{f_k\}$, $k = 1, 2, \ldots$, suite partout dense dans $\mathscr{C}_X$. Par la méthode de diagonalisation on peut choisir une suite de mesures μ_{n_k} telle que existe

$$\lim_{k\to\infty} \int f_i(x)\,\mu_{n_k}(dx) = l[f_i], \forall i.$$

Comme

$$|l[f_i] - l[f_j]| \leqslant H \|f_i - f_j\|,$$

alors $l[f_i]$ est une fonctionnelle uniformément continue sur l'ensemble $\{f_k\}$, $k = 1, 2, \ldots$, qui par conséquent est prolongeable par continuité à l'espace $\mathscr{C}_X$ tout entier: $l[f] = \lim_{f_{n_i}\to f} l[f_{n_i}]$. Ceci étant, la relation

$$l[f] = \lim_{k\to\infty} \int f(x)\,\mu_{n_k}(dx)$$

reste valable pour tous les f. La fonctionnelle $l[f]$ est linéaire non négative comme limite d'une suite de fonctionnelles linéaires non

négatives. Donc, en vertu du théorème de la forme d'une fonctionnelle linéaire sur l'espace $\mathscr{C}_X$ (cf. par exemple R. Edwards [1]), $l\,[f]$ s'écrit

$$l\,[f]=\int\limits_X f(x)\,\mu(dx),$$

où μ est une fonction d'ensemble dénombrablement additive non négative. Donc μ est une mesure et $\mu_{n_k}\Rightarrow\mu$. ■

Passons à la démonstration du théorème proprement dite.

Soient une suite $\varepsilon_m\to 0$ et des compacts $K^{(m)}\subset K^{(m+1)}$ tels que $\sup\limits_n \mu_n(X\setminus K^{(m)})\leqslant\varepsilon_m$. Posons

$$\mu_n^{(m)}(A)=\mu_n(A\cap K^{(m)}).$$

Prenons une suite $n_k^{(1)}$ telle que la suite de mesures $\mu^{(1)}_{n_k^{(1)}}$ converge faiblement vers une mesure $\mu^{(1)}$. Définissons une suite $n_k^{(j)}$ telle qu'elle soit sous-suite de $n_k^{(j-1)}$ et que $\mu^{(j)}_{n_k^{(j)}}$ soit faiblement convergente vers une mesure $\mu^{(j)}$. On a var $|\,\mu^{(j)}-\mu^{(j+p)}\,|\leqslant 2\,\varepsilon_j$, puisque $\mu^{(j)}$ et $\mu^{(j-1)}$ sont confondues sur $K^{(j-1)}$; donc la suite $\mu^{(j)}$ converge en variation vers une mesure μ. Montrons que $\mu_{n_k^{(k)}}$ converge faiblement vers μ. En effet, pour toute fonction continue on a

$$\varlimsup_{k\to\infty}\left|\int f(x)\,\mu_{n_k^{(k)}}(dx)-\int f(x)\,\mu(dx)\right|\leqslant$$
$$\leqslant\varlimsup_{k\to\infty}\left|\int\limits_{K^{(m)}} f(x)\,\mu_{n_k^{(k)}}(dx)-\int\limits_{K^{(m)}} f(x)\,\mu(dx)\right|+$$
$$+\,\|f\|\,(\varlimsup_{k\to\infty}\mu_{n_k^{(k)}}(X\setminus K^{(m)})+\mu(X\setminus K^{(m)}))\leqslant 2\,\|f\|\,\varepsilon_m.$$

D'où la condition suffisante.

Pour prouver la condition nécessaire on aura besoin du

LEMME 2. *On peut exhiber un compact* K *tel que* $\mu(X\setminus K)<\varepsilon$, $\forall\varepsilon>0$.

En effet, soient $\{x_k,\ k=1,2,\ldots\}$ suite partout dense dans X, S_k^n boule de rayon $1/2^n$ centrée en x_k. Comme

$$\bigcup_k S_k^n=X,$$

pour tout n existe k_n tel que

$$\mu\Big(\bigcup_{k=1}^{k_n} S_k^n\Big)\geqslant\mu(X)-\varepsilon/2^n.$$

En posant $K = \bigcap_{n=1}^{\infty} \bigcup_{k=1}^{kn} S_k^n$, on obtient un compact pour lequel

$$\mu(X \setminus K) \leqslant \sum_{n=1}^{\infty} \mu(X \setminus \bigcup_{k=1}^{kn} S_k^n) \leqslant \varepsilon \sum_{n=1}^{\infty} \frac{1}{2^n} = \varepsilon. \blacksquare$$

Condition nécessaire.

a) Si μ_n est faiblement compacte, alors $\int 1 \cdot \mu_n\,(dx)$ est un ensemble numérique compact et par conséquent la suite $\mu_n\,(X)$ est bornée.

Supposons maintenant que la suite μ_n est faiblement compacte mais que la condition b) n'est pas réalisée. On remarquera que la condition b) équivaut à la suivante:

b') $\forall \varepsilon > 0$, $\forall \delta > 0$ $\exists K$ compact tel que $\sup_n \mu_n\,(X \setminus K_\delta) < \varepsilon$, où K_δ désigne un ensemble de points x dont la distance à K est $\leqslant \delta$. Il est évident que b') implique b). Inversement, soit $K_{1/r}^{(r)}$ un compact tel que

$$\sup_n \mu_n\,(X \setminus K_{1/r}^{(r)}) \leqslant \varepsilon/2^r.$$

Alors $\bigcap_r K_{1/r}^{(r)}$ sera un compact vérifiant b). La non-réalisation de la condition b') équivaut à l'existence de $\varepsilon > 0$ et $\delta > 0$ tels que $\sup_n \mu_n\,(X \setminus K_\delta) > \varepsilon$, $\forall K$.

Appelons $K^{(0)}$ un compact tel que $\mu_1\,(X \setminus K^{(0)}) < \varepsilon$ (l'existence de $K^{(0)}$ est une conséquence du lemme 2). Comme $\sup_n \mu_n\,(X \setminus K_\delta^{(0)}) > \varepsilon$, il existe n_1 tel que

$$\mu_{n_1}(X \setminus K_\delta^{(0)}) > \varepsilon,$$

donc il existe un compact $K^{(1)}$ tel que $\mu_{n_1}(K^{(1)}) > \varepsilon$ et $K^{(1)} \subset X \setminus K_\delta^{(0)}$ (toujours en vertu du lemme 2). Comme

$$\sup_n \mu_n\,(X \setminus K_\delta^{(0)} \setminus K_\delta^{(1)}) > \varepsilon,$$

il existe n_2 et un compact $K^{(2)} \subset X \setminus K_\delta^{(0)} \setminus K_\delta^{(1)}$ tels que

$$\mu_{n_2}(K^{(2)}) > \varepsilon.$$

En poursuivant cette procédure on obtient une suite d'indices n_j et de compacts $K^{(j)}$ tels que

$$\mu_{n_j}(K^{(j)}) > \varepsilon \text{ et } K^{(j)} \subset X \setminus \bigcup_{i=0}^{j-1} K_\delta^{(i)} = X \setminus [\bigcup_{i=0}^{j-1} K^{(i)}]_\delta.$$

Soit $\chi_i\,(x)$ une fonction continue, non négative, bornée par l'unité, nulle sur $X \setminus K_{\delta/2}^{(i)}$ et égale à l'unité sur $K^{(i)}$. La distance entre

deux compacts quelconques de la suite $K^{(i)}$ étant $\geqslant \delta$, les fonctions $\chi_i(x)$ ne peuvent être simultanément non nulles. Considérons une sous-suite μ'_k faiblement convergente vers μ de la suite μ_{n_j}. La mesure μ étant finie et $\sum_i \chi_i(x)$ continue et bornée, il vient

$$\int \sum_{i=1}^{\infty} \chi_i(x)\,\mu(dx) = \sum_{i=1}^{\infty} \int \chi_i(x)\,\mu(dx) < \infty,$$

par suite

$$\lim_{p\to\infty} \sum_{i=p}^{\infty} \int \chi_i(x)\,\mu(dx) = 0. \text{ D'autre part}$$

$$\int \sum_{i=p}^{\infty} \chi_i(x)\,\mu'_k(dx) \geqslant \mu_{n_{p'}}(K^{(p')}) \geqslant \varepsilon \quad (\mu'_k = \mu_{n_{p'}}),$$

pourvu que $n_{p'} > p$ et par suite

$$\sum_{i=p}^{\infty} \int \chi_i(x)\,\mu(dx) = \lim_{k\to\infty} \sum_{i=p}^{\infty} \int \chi_i(x)\,\mu'_k(dx) \geqslant \varepsilon, \ \forall p.$$

Cette contradiction prouve la condition nécessaire b). ■

REMARQUE 1. *La complétude de l'espace X a servi à démontrer les conditions nécessaires du théorème. Les conditions du théorème sont des conditions suffisantes de faible compacité d'une suite de mesures dans tout espace métrique.*

REMARQUE 2. *Si μ_n converge faiblement vers μ, alors*

$$\lim_{n\to\infty} \mu_n(A) = \mu(A) \quad \forall A \in \mathfrak{B} \ \textit{tel que}\ \mu(A') = 0,$$

où A' est la frontière de A.

En effet, prenons un ensemble quelconque $A \in \mathfrak{B}$. Soient $A^{(0)}$ son noyau ouvert (c'est-à-dire l'ensemble de tous ses points intérieurs), $[A]$ son adhérence. Si μ_n converge faiblement vers μ, en prenant une fonction $f(x) \geqslant 0$ continue telle que $f(x) = 1$ pour $x \in [A]$ et

$$\mu([A]) \geqslant \int f(x)\,\mu(dx) - \varepsilon,$$

on trouve

$$\mu([A]) \geqslant \int f(x)\,\mu(dx) - \varepsilon = \lim \int f(x)\,\mu_n(dx) - \varepsilon \geqslant \overline{\lim}\,\mu_n(A) - \varepsilon.$$

Donc

$$\overline{\lim}\, \mu_n (A) \leqslant \mu ([A]),\ \overline{\lim}\, \mu_n (X \setminus A) \leqslant \mu ([X \setminus A]),$$
$$-\underline{\lim}\, \mu_n (A) \leqslant -\mu (A^{(0)}).$$

Par suite

$$\mu (A^{(0)}) \leqslant \underline{\lim}\, \mu_n (A) \leqslant \overline{\lim}\, \mu_n (A) \leqslant \mu ([A]).$$

Si $\mu (A') = 0$, alors $\mu (A^{(0)}) = \mu ([A]) = \mu (A)$ et

$$\mu (A) \leqslant \varliminf_{n\to\infty} \mu_n (A) \leqslant \varlimsup_{n\to\infty} \mu_n (A) \leqslant \mu (A).$$

D'où suit l'assertion de la remarque.

REMARQUE 3. *Si $f(x)$ est presque partout continue par rapport à la mesure μ et μ_n converge faiblement vers μ, alors*

$$\lim_{n\to\infty} \mu_n (\{x : f(x) < \alpha\}) = \mu (\{x : f(x) < \alpha\})$$

pour presque tous les α.

En effet, soit A_0 l'ensemble des points de discontinuité de f. On a $\mu (A_0) = 0$. Soient G_α l'ensemble des x tels que $f(x) < \alpha$ et G'_α la frontière de G_α :

$$G'_\alpha = [\{x : f(x) < \alpha\}] \cap [\{x : f(x) \geqslant \alpha\}].$$

L'intersection des ensembles G'_α et G'_{α_1}, $\alpha < \alpha_1$, est contenue dans celle des ensembles $[\{x : f(x) < \alpha\}] \cap [\{x : f(x) \geqslant \alpha_1\}]$; donc $x \in G'_\alpha \cap G'_{\alpha_1}$ implique

$$\lim_{y\to x} \inf f(y) \leqslant \alpha, \qquad \lim_{y\to x} \sup f(y) \geqslant \alpha_1,$$

c'est-à-dire

$$G'_\alpha \cap G'_{\alpha_1} \subset A_0.$$

Donc $\mu (G'_\alpha \cap G'_{\alpha_1}) = 0$ et

$$\mu (\bigcup_k G'_{\alpha_k}) = \sum_k \mu (G'_{\alpha_k})$$

pour toute suite α_k.

D'où il suit qu'existe un ensemble au plus dénombrable de nombres α tels que $\mu (G'_\alpha) \neq 0$.

Donc pour tous les α sauf peut-être pour un nombre dénombrable d'entre eux, G_α est ensemble de continuité de mesure μ, de sorte que $\mu_n (G_\alpha) \to \mu (G_\alpha)$.

Ceci prouve notre assertion.

REMARQUE 4. Pour démontrer les théorèmes limites pour processus aléatoires on se sert du théorème 1 de la manière suivante. Soit

$\xi_n(t)$ suite de processus aléatoires dont les répartitions finidimensionnelles convergent vers celles d'un processus $\xi_0(t)$. Si μ_n, mesures associées aux $\xi_n(t)$, forment un ensemble compact, alors μ_n convergent faiblement vers μ_0, mesure associée au processus $\xi_0(t)$. En effet, dans le cas contraire on exhiberait une sous-suite n_k telle que μ_{n_k} convergent faiblement vers une mesure $\overline{\mu} \neq \mu_0$. Soit $\overline{\xi}(t)$ le processus auquel est associée la mesure $\overline{\mu}$. Les répartitions finidimensionnelles de $\overline{\xi}(t)$, comme limite des répartitions finidimensionnelles de $\xi_{n_k}(t)$, sont confondues avec celles de $\xi_0(t)$, ce qui n'est possible que pour $\overline{\mu} = \mu_0$, car les mesures associées aux processus sont définies de façon unique par leurs répartitions finidimensionnelles.

§ 2. Théorèmes limites pour processus continus

Dans ce paragraphe on supposera que les processus $\xi_n(t)$ et $\xi(t)$ sont continus sur l'intervalle $[a, b]$. Leurs réalisations appartiennent presque sûrement à l'espace $\mathscr{C}_{[a,b]}$] des fonctions $x(t)$ continues sur $[a, b]$, qui est un espace polonais muni de la métrique $\rho(x, y) = \sup\limits_{a \leqslant t \leqslant b} |x(t) - y(t)|$.

A noter que la plus petite tribu $\mathfrak{A}$ de $\mathscr{C}_{[a,b]}$ contient tous les ensembles cylindriques et tous les boréliens. Pour le vérifier il suffit de remarquer que toute boule fermée appartient à $\mathfrak{A}$, puisque

$$\{x : \sup_t |x(t) - \alpha(t)| \leqslant r\} = \bigcap_{k=1}^{\infty} \{x : |x(t_k) - \alpha(t_k)| \leqslant r\},$$

où $\alpha(t)$ est une fonction continue quelconque et t_k une suite quelconque partout dense dans $[a, b]$.

Soient H une constante, ω_δ une fonction définie pour $\delta > 0$ telle que $\omega_\delta \downarrow 0$ pour $\delta \downarrow 0$. Désignons par $K(H, \omega_\delta)$ l'ensemble des fonctions $x(t)$ telles que

$$x(a) \leqslant H, \ \forall \delta > 0 \text{ et } \sup_{|t'-t''| \leqslant \delta} |x(t') - x(t'')| \leqslant \omega_\delta.$$

Le théorème d'Arzelà (cf. A. Kolmogorov et S. Fomine [1], page 104) nous apprend que tout compact de $\mathscr{C}_{[a,b]}$ est partie fermée d'un ensemble $K(H, \omega_\delta)$ également compact.

THÉORÈME 1. *Supposons que les répartitions finidimensionnelles de processus $\xi_n(t)$ convergent vers les répartitions finidimensionnelles d'un processus $\xi(t)$. Une condition nécessaire et suffisante pour que la répartition de $f(\xi_n(t))$ converge vers celle de $f(\xi(t))$ $\forall f \in \mathscr{C}_{[a,b]}$ est que*

$$\lim_{h \to 0} \sup_n \mathsf{P}\{\sup_{|t'-t''| \leqslant h} |\xi_n(t') - \xi_n(t'')| > \varepsilon\} = 0, \ \forall \varepsilon > 0. \qquad (1)$$

D é m o n s t r a t i o n. *Condition nécessaire.* Si le théorème est vrai, la suite de mesures μ_n associées aux processus $\xi_n(t)$ est faiblement compacte, de sorte qu'est réalisée la condition b) du théorème 1, § 1. Donc pour tout $\eta > 0$ il existe un compact $K(H, \omega_\delta)$ tel que

$$\sup_n \mu_n(\mathscr{C}_{[a,b]} \setminus K(H, \omega_\delta)) = \sup_n \mathsf{P}\{\xi_n(t) \notin K(H, \omega_\delta)\} \leqslant \eta.$$

On a

$$\mathsf{P}\{\sup_{|t'-t''| \leqslant h} |\xi_n(t') - \xi_n(t'')| > \omega_h\} \leqslant \mathsf{P}\{\xi_n(t) \notin K(H, \omega_h)\} \leqslant \eta.$$

Si h est assez petit, alors $\omega_h < \varepsilon$ et

$$\overline{\lim_{h \to 0}} \sup_n \mathsf{P}\{|\xi_n(t') - \xi_n(t'')| > \varepsilon\} \leqslant \eta.$$

D'où l'on déduit (1) puisque $\eta > 0$ est arbitraire.

Condition suffisante. La remarque 4, § 1, indique qu'il suffit de prouver la faible compacité de la suite de mesures μ_n. Montrons que pour tout $\eta > 0$ il existe un compact $K(H, \omega_\delta)$ tel que

$$\sup_n \mathsf{P}\{\xi_n(t) \notin K(H, \omega_\delta)\} \leqslant \eta.$$

La répartition de $\xi_n(a)$ convergeant vers celle de $\xi(a)$, il existe H tel que

$$\mathsf{P}\{|\xi_n(a)| > H\} \leqslant \eta/2, \ \forall n.$$

Considérons une suite $\varepsilon_r \downarrow 0$. Pour tout ε_r il existe h_r tel que $h_r < h_{r-1}$ et

$$\sup_n \mathsf{P}\{\sup_{|t'-t''| < h_r} |\xi_n(t') - \xi_n(t'')| > \varepsilon_r\} \leqslant \frac{\eta}{2^{r+1}}.$$

Soit ω_δ une fonction décroissante non négative telle que $\omega_\delta = \varepsilon_r$ pour $\delta \in [h_{r+1}, h_r[$. Il est évident que $\omega_\delta \downarrow 0$ pour $\delta \downarrow 0$. D'autre part,

$$\mathsf{P}\{\xi_n(t) \notin K(H, \omega_\delta)\} \leqslant \mathsf{P}\{|\xi_n(a)| > H\} + \sum_{r=1}^{\infty} \mathsf{P}\{\sup_{|t'-t''| \leqslant h_r} |\xi_n(t') - \xi_n(t'')| > \varepsilon_r\} \leqslant \frac{\eta}{2} + \sum_{r=1}^{\infty} \frac{\eta}{2^{r+1}} = \eta. \ \blacksquare$$

Remarque 1. Au lieu de (1) on peut se servir de la condition

$$\lim_{h \to 0} \overline{\lim_{n \to \infty}} \mathsf{P}\{\sup_{|t'-t''| \leqslant h} |\xi_n(t') - \xi_n(t'')| > \varepsilon\} = 0, \tag{2}$$

qui souvent est plus facile à vérifier.

En effet, de (2) il suit que pour tout $\eta > 0$ il existe $\delta > 0$ et N tels que

$$\mathsf{P}\{\sup_{|t'-t''| \leqslant h} |\xi_n(t') - \xi_n(t'')| > \varepsilon\} \leqslant \eta \tag{3}$$

pour $n > N$, $h < \delta$.

Etant continus, les processus $\xi_n(t)$ le seront uniformément de sorte que pour tout n

$$\lim_{h\to 0} \mathsf{P}\{\sup_{|t'-t''|\leqslant h} |\xi_n(t')-\xi_n(t'')|>\varepsilon\}=0.$$

On peut donc choisir un δ tel que pour $h<\delta$ la relation (3) soit réalisée pour tous les n.

Le théorème suivant est susceptible d'être plus commode à l'usage.

THÉORÈME 2. *Si les répartitions finidimensionnelles de processus* $\xi_n(t)$ *tendent vers celles d'un processus* $\xi(t)$ *et qu'existent* $\alpha>0$, $\beta>0$ *et* $H>0$ *tels que pour tous les* t_1, t_2 *et tous les* n

$$\mathsf{E}\,|\xi_n(t_1)-\xi_n(t_2)|^\alpha\leqslant H\,|t_1-t_2|^{1+\beta}, \tag{4}$$

alors la répartition de $f(\xi_n(t))$ *convergera vers celle de* $f(\xi(t))$ $\forall f\in \mathscr{C}_{[a,b]}$.

D é m o n s t r a t i o n. Les processus $\xi_n(t)$ étant continus, on a

$$\sup_{|t_1-t_2|\leqslant h} |\xi_n(t_1)-\xi_n(t_2)| = \sup_{\substack{|t_1-t_2|\leqslant h\\ t_1,\, t_2\in N}} |\xi_n(t_1)-\xi_n(t_2)|,$$

où N est l'ensemble de tous les points de la forme $k/2^m$ de $[a, b]$. Si $2h<1/2^k$, alors

$$\sup_{\substack{t',\,t''\in N\\ |t'-t''|\leqslant h}} |\xi_n(t')-\xi_n(t'')|\leqslant 2\sup_j \sup_{\frac{j}{2^k}<\frac{l}{2^m}<\frac{j+1}{2^k}} \left|\xi_n\left(\frac{l}{2^m}\right)-\xi_n\left(\frac{j}{2^k}\right)\right|\leqslant$$

$$\leqslant 2\sum_{m=k+1}^{\infty} \sup_l \left|\xi_n\left(\frac{l+1}{2^m}\right)-\xi_n\left(\frac{l}{2^m}\right)\right|,$$

puisque $\frac{l}{2^m}-\frac{j}{2^k}=\sum_{r=1}^{s}\frac{1}{2^{m_r}}$, où $k<m_1<m_2<\ldots<m_s\leqslant m$, et partant

$$\xi_n\left(\frac{l}{2^m}\right)-\xi_n\left(\frac{j}{2^k}\right)=\sum_{i=1}^{s}\left[\xi_n\left(\frac{j}{2^k}+\sum_{r=1}^{i}\frac{1}{2^{m_r}}\right)-\xi_n\left(\frac{j}{2^k}+\sum_{r=1}^{i-1}\frac{1}{2^{m_r}}\right)\right].$$

On remarquera que

$$\mathsf{P}\left\{\sup_{a\leqslant\frac{i}{2^m}<\frac{i+1}{2^m}\leqslant b}\left|\xi_n\left(\frac{i+1}{2^m}\right)-\xi_n\left(\frac{i}{2^m}\right)\right|>\frac{1}{m^2}\right\}\leqslant$$

$$\leqslant\sum_i \mathsf{P}\left\{\left|\xi_n\left(\frac{i+1}{2^m}\right)-\xi_n\left(\frac{i}{2^m}\right)\right|>\frac{1}{m^2}\right\}\leqslant$$

$$\leqslant m^{2\alpha} \sum_i \mathsf{E} \left| \xi_n \left(\frac{i+1}{2^m} \right) - \xi_n \left(\frac{i}{2^m} \right) \right|^\alpha \leqslant$$

$$\leqslant m^{2\alpha} (b-a)\, 2^m H \frac{1}{2^{m(1+\beta)}} = L \frac{m^{2\alpha}}{2^{m\beta}}.$$

Donc, si

$$\sum_{m=k+1}^{\infty} \frac{1}{m^2} < \frac{\varepsilon}{2}, \quad k = \left[\log_2 \frac{1}{h} \right] + 1,$$

alors

$$\mathsf{P}\{ \sup_{|t_1-t_2|\leqslant h} |\xi_n(t_1) - \xi_n(t_2)| > \varepsilon \} \leqslant$$

$$\leqslant \sum_{m=k+1}^{\infty} \mathsf{P} \left\{ \sup_i \left| \xi_n \left(\frac{i+1}{2^m} \right) - \xi_n \left(\frac{i}{2^m} \right) \right| > \frac{1}{m^2} \right\} \leqslant$$

$$\leqslant L \sum_{m \geqslant \left[\log_2 \frac{1}{h}\right]+2} \frac{m^{2\alpha}}{2^{m\beta}} \to 0$$

uniformément en n pour $h \to 0$. ■

§ 3. Convergence de sommes de variables aléatoires indépendantes vers un processus de mouvement brownien

Considérons une suite de séries de variables aléatoires $\xi_{n1}, \xi_{n2}, \ldots$ $\ldots, \xi_{nk_n}$, indépendantes dans chaque série et telles que:

1) $\mathsf{E}\xi_{ni} = 0$;

2) $\operatorname{Var} \xi_{ni} = b_{ni}$, $\sum_{i=1}^{k_n} b_{ni} = 1$.

Construisons une fonction aléatoire $\xi_n(t)$, $t \in [0, 1]$ de la façon suivante: posons $S_{nk} = \sum_{i=1}^{k} \xi_{ni}$, $t_{nk} = \sum_{i=1}^{k} b_{ni}$,

$$\xi_n(t) = S_{nk} + \frac{t - t_{nk}}{t_{n,\,k+1} - t_{nk}} [S_{n,\,k+1} - S_{nk}]$$

pour $t \in [t_{nk}, t_{n,\,k+1}]$, $S_{n0} = 0$, $t_{n0} = 0$. Le graphe de la fonction $\xi_n(t)$ est une ligne polygonale aléatoire reliant les points du plan (t, ξ) de coordonnées (t_{nk}, S_{nk}), $k = 0, 1, \ldots, k_n$.

Dans ce paragraphe on se propose d'étudier les conditions sous lesquelles les répartitions finidimensionnelles des processus $\xi_n(t)$ et les répartitions de fonctionnelles de ces processus convergent vers

les répartitions finidimensionnelles et les répartitions des fonctionnelles respectives d'un processus de mouvement brownien $w(t)$.

THÉORÈME 1. *Supposons que les variables aléatoires* ξ_{ni} *vérifient les conditions* 1) *et* 2) *et la condition suivante de Lindeberg : si* $F_{ni}(x)$ *est fonction de répartition de* ξ_{ni}, *pour tout* ε

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} \int_{|u|>\varepsilon} u^2\, dF_{ni}(u) = 0. \tag{1}$$

Alors les répartitions finidimensionnelles des processus $\xi_n(t)$ *convergent vers les répartitions finidimensionnelles d'un processus* $w(t)$ *et la répartition de* $f(\xi_n(t))$ *vers la répartition de* $f(w(t))$ *pour tout* $f \in \mathscr{C}_{[0,1]}$.

Démonstration. La convergence des répartitions finidimensionnelles des processus $\xi_n(t)$ vers celles d'un processus $w(t)$ est une conséquence du théorème limite central en dimension n.

Pour démontrer la convergence des répartitions de $f(\xi_n(t))$ vers celles de $f(w(t))$, $\forall f \in \mathscr{C}_{[0,1]}$, suivant la remarque 1, § 2, il nous faudra vérifier que

$$\lim_{h\to 0} \overline{\lim_{n\to\infty}} \mathsf{P}\{\sup_{|t'-t''|\leqslant h} |\xi_n(t')-\xi_n(t'')| > \varepsilon\} = 0, \quad \forall \varepsilon > 0. \tag{2}$$

Comme

$$\sup_{|t'-t''|\leqslant h} |\xi_n(t')-\xi_n(t'')| \leqslant 2 \sup_k \sup_{kh<t\leqslant(k+2)h} |\xi_n(t)-\xi_n(kh)| \leqslant$$
$$\leqslant 4 \sup_k \sup_{kh<t\leqslant(k+1)h} |\xi_n(t)-\xi_n(kh)|,$$

il vient

$$\mathsf{P}\{\sup_{|t'-t''|\leqslant h} |\xi_n(t')-\xi_n(t'')| > \varepsilon\} \leqslant$$
$$\leqslant \sum_{kh<1} \mathsf{P}\{\sup_{kh<t\leqslant(k+1)h} |\xi_n(t)-\xi_n(kh)| > \varepsilon/4\}.$$

A noter que

$$\sup_{kh<t\leqslant(k+1)h} |\xi_n(t)-\xi_n(kh)| \leqslant 2 \sup_{j_{nk}<r\leqslant j_{n,k+1}} \left| \sum_{j=j_{nk}}^{r} \xi_{nj} \right|,$$

où j_{nk} est le plus grand des indices j tels que $t_{nj} \leqslant kh$. Comme, pour $j_{nk} \leqslant r \leqslant j_{n,k+1}$

$$\mathsf{P}\left\{\left| \sum_{j=r}^{j_{n,k+1}} \xi_{nj} \right| > \frac{\varepsilon}{8}\right\} \leqslant \frac{64}{\varepsilon^2} \sum \operatorname{Var} \xi_{nj} \to \frac{64h}{\varepsilon^2},$$

le lemme 3, § 5, chapitre VI, donne pour les h assez petits

$$\overline{\lim_{n\to\infty}} \mathsf{P}\{\sup_{kh<t\leqslant(k+1)h} |\xi_n(t)-\xi_n(kh)|>\varepsilon/4\}\leqslant$$

$$\leqslant\frac{1}{1-\frac{64h}{\varepsilon^2}}\overline{\lim_{n\to\infty}} \mathsf{P}\{|\xi_n(t_{nj_{n,k+1}})-\xi_n(t_{nj_{nk}})|>\varepsilon/8\}.$$

La convergence des répartitions finidimensionnelles de $\xi_n(t)$ vers celles de $w(t)$ implique

$$\overline{\lim_{n\to\infty}} \mathsf{P}\{|\xi_n(t_{nj_{n,k+1}})-\xi_n(t_{nj_{nk}})|>\varepsilon/8\}=\frac{1}{\sqrt{2\pi h}}\int\limits_{|u|>\varepsilon/8} e^{-u^2/2h}\,du.$$

Donc

$$\overline{\lim_{n\to\infty}} \mathsf{P}\{\sup_{|t'-t''|\leqslant h} |\xi_n(t')-\xi_n(t'')|>\varepsilon\}\leqslant$$

$$\leqslant\sum_{kh<1}\frac{1}{1-\frac{64h}{\varepsilon^2}}\frac{1}{\sqrt{2\pi}}\int\limits_{|u|>\frac{\varepsilon}{8\sqrt{h}}} e^{-\frac{u^2}{2}}\,du\leqslant$$

$$\leqslant\frac{1}{\sqrt{2\pi}}\frac{1}{1-\frac{64h}{\varepsilon^2}}\frac{1}{h}\int\limits_{|u|>\frac{\varepsilon}{8\sqrt{h}}} e^{-\frac{u^2}{2}}\,du.$$

De la relation

$$\frac{1}{h}\int\limits_{|u|>\frac{c}{\sqrt{h}}} e^{-\frac{u^2}{2}}\,du\leqslant\frac{1}{c^2}\int\limits_{|u|>\frac{c}{\sqrt{h}}} u^2e^{-\frac{u^2}{2}}\,du\to 0$$

on déduit (2). ■

COROLLAIRE. *Soit* $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *suite de variables aléatoires indépendantes équiréparties telles que* $\mathsf{E}\xi_i = 0$, $\mathrm{Var}\,\xi_i = 1$.

Désignons par $\xi_n(t)$ *la ligne polygonale aléatoire de sommets* $\left(\frac{k}{n}, \frac{1}{\sqrt{n}} S_k\right)$, *où* $S_k = \xi_1 + \ldots + \xi_k$. *Alors la répartition de* $f(\xi_n(t))$ *convergera vers celle de* $f(w(t))$ *pour toute fonctionnelle définie et continue sur* $\mathscr{C}_{[0,1]}$ *presque partout par rapport à la mesure* μ_w.

§ 4. Convergence d'une suite de chaînes de Markov vers un processus de diffusion

Considérons une suite de séries de variables aléatoires $\xi_{n0}, \xi_{n1}, \ldots, \xi_{nk_n}$, formant une chaîne de Markov dans chaque série. Appelons $p_{nk}(x, A)$ les probabilités de passage

$$p_{nk}(\xi_{nk}, A)=\mathsf{P}\{\xi_{n,k+1}\in A\,|\,\xi_{nk}\}\quad (\mathrm{mod}\,\mathsf{P}).$$

Supposons ensuite que $0 = t_{n0} < t_{n1} < \ldots < t_{nk_n} = 1$ est une suite de partitions de l'intervalle [0, 1]. Construisons une ligne polygonale aléatoire $\xi_n(t)$ de sommets (t_{nk}, ξ_{nk}). Dans ce paragraphe on se propose d'étudier les critères de convergence des répartitions finidimensionnelles de $\xi_n(t)$ et des fonctionnelles de $\xi_n(t)$ vers les répartitions correspondantes d'un processus markovien $\xi(t)$, solution d'une équation différentielle stochastique du type étudié au chapitre VIII.

Posons

$$\Delta t_{nk} = t_{n,\,k+1} - t_{nk},$$

$$a_n(t_{nk}, x) = \frac{1}{\Delta t_{nk}} \int (y-x)\, p_{nk}(x, dy),$$

$$b_n(t_{nk}, x) = \sigma_n^2(t_{nk}, x) = \frac{1}{\Delta t_{nk}} \int (y-x)^2 p_{nk}(x, dy) - \\ - \Delta t_{nk} a_n^2(t_{nk}, x).$$

THÉORÈME 1. *Supposons que $\xi(t)$ est solution de l'équation différentielle stochastique*

$$\xi(t) = \xi_0 + \int_0^t a(s, \xi(s))\, ds + \int_0^t \sigma(s, \xi(s))\, dw(s),$$

où ξ_0 ne dépend pas de $w(t)$, et $a(s, x)$ et $\sigma(s, x)$ sont des fonctions continues en s et x et vérifiant une condition de Lipschitz par rapport à x: $|a(s, x) - a(s, y)| + |\sigma(s, x) - \sigma(s, y)| \leqslant K|x-y|$. Pour que les répartitions finidimensionnelles des processus $\xi_n(t)$ convergent vers celles du processus $\xi(t)$, il suffit que soient réalisées les conditions suivantes:

1) $\lim\limits_{n\to\infty} \max\limits_k \Delta t_{nk} = 0$;

2) $$\lim_{n\to\infty} \sum_{k=1}^{k_n} \mathsf{E}\,(|a_n(t_{nk}, \xi_{nk}) - a(t_{nk}, \xi_{nk})|^2 + \\ + |\sigma_n(t_{nk}, \xi_{nk}) - \sigma_n(t_{nk}, \xi_{nk})|^2)\, \Delta t_{nk} = 0;$$

3) *pour un* $\delta > 0$

$$\lim_{n\to\infty} \sum_{k=0}^{k_n-1} \mathsf{E}\,|\xi_{n,\,k+1} - \xi_{nk}|^{2+\delta} = 0$$

et

$$\mathsf{P}\text{-}\lim_{n\to\infty} \sup_{k\geqslant 0} \sum_{i=k}^{k_n-1} \mathsf{E}\,(|\xi_{n,\,i+1} - \xi_{ni}|^{2+\delta} \mid \xi_{nk}) = 0;$$

4) *les fonctions* $\frac{1}{\sigma_n(t_{nk}, x)}$ *et* $\frac{a_n(t_{nk}, x)}{\sigma_n(t_{nk}, x)}$ *sont uniformément bornées par rapport à* n;

5) *la répartition limite de la variable* ξ_{n0} *coïncide avec celle de* ξ_0.

Démonstration. Posons

$$\omega_{nk} = [\xi_{n,k+1} - \xi_{nk} - a_n(t_{nk}, \xi_{nk})\Delta t_{nk}](\sigma_n(t_{nk}, \xi_{nk}))^{-1}. \qquad (1)$$

Alors

$$\xi_{n,k+1} = \xi_{nk} + a_n(t_{nk}, \xi_{nk})\Delta t_{nk} + \sigma_n(t_{nk}, \xi_{nk})\omega_{nk}.$$

Appelons $\mathfrak{F}_{nk}$ la plus petite tribu par rapport à laquelle sont mesurables $\xi_{n0}, \xi_{n1}, \ldots, \xi_{nk}$. La variable ω_{nk} est $\mathfrak{F}_{n,k+1}$-mesurable et de plus

$$\mathsf{E}(\omega_{nk} \mid \mathfrak{F}_{nk}) = 0, \qquad \mathsf{E}(\omega_{nk}^2 \mid \mathfrak{F}_{nk}) = \Delta t_{nk}. \qquad (2)$$

Soient

$$\eta_{n0} = \xi_{n0},$$
$$\eta_{n,k+1} = \eta_{nk} + a(t_{nk}, \eta_{nk})\Delta t_{nk} + \sigma(t_{nk}, \eta_{nk})\omega_{nk}.$$

Evaluons $\mathsf{E}(\eta_{nk} - \xi_{nk})^2$. De toute évidence, η_{nk} sont $\mathfrak{F}_{nk}$-mesurables. On a

$$\eta_{n,k+1} - \xi_{n,k+1} = \eta_{nk} - \xi_{nk} + [a(t_{nk}, \eta_{nk}) - a(t_{nk}, \xi_{nk})]\Delta t_{nk} + \\ + [\sigma(t_{nk}, \eta_{nk}) - \sigma(t_{nk}, \xi_{nk})]\omega_{nk} + \varepsilon_{nk},$$

$$\varepsilon_{nk} = [a(t_{nk}, \xi_{nk}) - a_n(t_{nk}, \xi_{nk})]\Delta t_{nk} + \\ + [\sigma(t_{nk}, \xi_{nk}) - \sigma_n(t_{nk}, \xi_{nk})]\omega_{nk}.$$

Les relations (2), les conditions de Lipschitz pour a et σ et l'inégalité $2ab \leqslant a^2 + b^2$ donnent

$$\mathsf{E}|\eta_{n,k+1} - \xi_{n,k+1}|^2 \leqslant \\ \leqslant \mathsf{E}|\eta_{nk} - \xi_{nk}|^2 + 2\mathsf{E}(\eta_{nk} - \xi_{nk})(a(t_{nk}, \eta_{nk}) - a(t_{nk}, \xi_{nk}))\Delta t_{nk} + \\ + 2\mathsf{E}(\eta_{nk} - \xi_{nk})(a(t_{nk}, \xi_{nk}) - a_n(t_{nk}, \xi_{nk}))\Delta t_{nk} + \\ + \mathsf{E}[(a(t_{nk}, \eta_{nk}) - a(t_{nk}, \xi_{nk}))\Delta t_{nk} + (\sigma(t_{nk}, \eta_{nk}) - \\ - \sigma(t_{nk}, \xi_{nk}))\omega_{nk} + \varepsilon_{nk}]^2 \leqslant \mathsf{E}|\eta_{nk} - \xi_{nk}|^2(1 + 2K\Delta t_{nk} + \Delta t_{nk}) + \\ + \mathsf{E}|a(t_{nk}, \xi_{nk}) - a_n(t_{nk}, \xi_{nk})|^2\Delta t_{nk} + \\ + 2\mathsf{E}(a(t_{nk}, \eta_{nk}) - a(t_{nk}, \xi_{nk}))^2\Delta t_{nk}^2 + \\ + \mathsf{E}(\sigma(t_{nk}, \eta_{nk}) - \sigma(t_{nk}, \xi_{nk}))^2\mathsf{E}(\omega_{nk}^2 \mid \mathfrak{F}_{nk}) + 2\mathsf{E}\varepsilon_{nk}^2 \leqslant \\ \leqslant \mathsf{E}|\eta_{nk} - \xi_{nk}|^2(1 + L\Delta t_{nk}) + \alpha_{nk},$$

où $L = 2K + 1 + 4K^2$,

$$\alpha_{nk} = \mathsf{E}[a(t_{nk}, \xi_{nk}) - a_n(t_{nk}, \xi_{nk})]^2(\Delta t_{nk} + 2\Delta t_{nk}^2) + \\ + \mathsf{E}[\sigma(t_{nk}, \xi_{nk}) - \sigma_n(t_{nk}, \xi_{nk})]^2\Delta t_{nk}.$$

Comme

$$\mathsf{E}\,|\,\eta_{n0}-\xi_{n0}\,|^2=0,$$

on trouve

$$\mathsf{E}\,|\,\xi_{n1}-\eta_{n1}\,|^2\leqslant\alpha_{n0},$$

$$\mathsf{E}\,|\,\xi_{n2}-\eta_{n2}\,|^2\leqslant\alpha_{n0}\,(1+L\,\Delta t_{n0})+\alpha_{n1}\leqslant(1+L\,\Delta t_{n0})\,[\alpha_{n0}+\alpha_{n1}],$$

$$\mathsf{E}\,|\,\xi_{n3}-\eta_{n3}\,|^2\leqslant[\alpha_{n0}+\alpha_{n1}]\,(1+L\,\Delta t_{n0})\,(1+L\,\Delta t_{n1})+\alpha_{n2}\leqslant$$

$$\leqslant[\alpha_{n0}+\alpha_{n1}+\alpha_{n2}]\,(1+L\,\Delta t_{n0})\,(1+L\,\Delta t_{n1}),$$

$$\mathsf{E}\,|\,\xi_{nk}-\eta_{nk}\,|^2\leqslant\sum_{i=0}^{k-2}\alpha_{ni}\prod_{i=0}^{k-2}(1+L\,\Delta t_{ni})\leqslant e^{L}\sum_{i=0}^{k_n-1}\alpha_{ni}.$$

La condition 2) entraîne que $\lim\limits_{n\to\infty}\sum\limits_{i=0}^{k_n-1}\alpha_{ni}=0.$

Donc les répartitions finidimensionnelles du processus $\xi_n(t)$ convergeront vers celles du processus $\xi(t)$ si seulement les répartitions finidimensionnelles de $\xi(t)$ seront limite de celles du processus $\eta_n(t)$, où $\eta_n(t)$ est une ligne polygonale de sommets (t_{nk}, η_{nk}). Pour achever la démonstration nous aurons besoin du

LEMME 1. *Si* $w_n(t)=\sum\limits_{t_{nk}<t}\omega_{nk}$, *alors les répartitions finidimensionnelles du processus* $w_n(t)$ *convergeront vers celles d'un processus wienérien* $w(t)$.

Il suffit de montrer que pour tous les λ_{nk} tels que $\sup\limits_{n,k}|\,\lambda_{nk}\,|<\infty$, on aura

$$\lim_{n\to\infty}\Big[\mathsf{E}\exp\Big\{i\sum_{k=0}^{k_n-1}\lambda_{nk}\omega_{nk}\Big\}-$$

$$-\mathsf{E}\exp\Big\{i\sum_{k=0}^{k_n-1}\lambda_{nk}\,[w(t_{n,\,k+1})-w(t_{nk})]\Big\}\Big]=0.$$

On remarquera que la fonction $\dfrac{e^{i\lambda x}-1-i\lambda x+\dfrac{\lambda^2x^2}{2}}{|\,\lambda\,|^2\,|\,x\,|^{2+\delta}}$ est bornée sur la droite numérique tout entière pour $0<\delta<1$. Donc

$$\Big|\,e^{i\lambda x}-\Big(1+i\lambda x-\frac{\lambda^2x^2}{2}\Big)\Big|\leqslant O(\lambda^2)\,|\,x\,|^{2+\delta}$$

et

$$\mathsf{E}\prod_{k=1}^{r} e^{i\lambda_{nk}\omega_{nk}} = \mathsf{E}\left(\prod_{k=0}^{r-1} e^{i\lambda_{nk}\omega_{nk}}\right)\mathsf{E}\left[e^{i\lambda_{nr}\omega_{nr}} \mid \mathfrak{F}_{nr}\right] =$$

$$= \mathsf{E}\left(\prod_{k=0}^{r-1} e^{i\lambda_{nk}\omega_{nk}}\right)\mathsf{E}\left[1+\lambda_{nr}\omega_{nr} - \frac{\lambda_{nr}^2}{2}\omega_{nr}^2 + O(|\omega_{nr}|^{2+\delta}) | \mathfrak{F}_{nr}\right] =$$

$$= \mathsf{E}\left(\prod_{k=0}^{r-1} e^{i\lambda_{nk}\omega_{nk}}\right)\left(1-\frac{\lambda_{nr}^2}{2}\Delta t_{nr}\right) + O(\mathsf{E}|\omega_{nr}|^{2+\delta}) =$$

$$= \prod_{k=0}^{r}\left(1-\frac{\lambda_{nk}^2}{2}\Delta t_{nk}\right) + O\left(\sum_{k=0}^{r}\mathsf{E}|\omega_{nk}|^{2+\delta}\right).$$

L'inégalité

$$\mathsf{E}|\omega_{nk}|^{2+\delta} \leqslant L(\mathsf{E}|\xi_{n,k+1} - \xi_{nk}|^{2+\delta} + |\Delta t_{nk}|^{2+\delta})$$

découle de la formule (1) pour ω_{nk} et fait que $\frac{1}{\sigma_n}$ et $\frac{a_n}{\sigma_n}$ sont bornés.

Donc

$$\lim_{n\to\infty}\sum_{k=0}^{k_n-1}\mathsf{E}|\omega_{nk}|^{2+\delta} = 0.$$

D'autre part

$$\left|\prod_{k=0}^{k_n-1}\left(1-\frac{\lambda_{nk}^2}{2}\Delta t_{nk}\right) - \prod_{k=0}^{k_n-1} e^{-\frac{\lambda_{nk}^2}{2}\Delta t_{nk}}\right| \leqslant O\left(\sum \Delta t_{nk}^2\right) =$$

$$= O(\max_k \Delta t_{nk}) \to 0.$$

Par suite

$$\lim_{n\to\infty}\left(\mathsf{E}e^{i\sum_{k=0}^{k_n-1}\lambda_{nk}\omega_{nk}} - e^{-\sum_{k=0}^{k_n-1}\frac{\lambda_{nk}^2}{2}\Delta t_{nk}}\right) = 0.$$

Reste à remarquer que

$$\mathsf{E}\exp\{i\lambda_{nk}[w(t_{n,k+1}) - w(t_{nk})]\} = e^{-\frac{\lambda_{nk}^2}{2}\Delta t_{nk}}. \blacksquare$$

REMARQUE. *Du lemme précédent il suit que pour toute fonction $\alpha(t, x)$ continue*

$$\mathsf{E}\left\{\exp\left(i\lambda\sum_{t'\leqslant t_{nk}\leqslant t''}\alpha(t_{nk}, x)\,\omega_{nk}\right)\Big|\mathfrak{F}_{t_n}\right\} \to$$

$$\to \mathsf{E}\left\{\exp i\lambda \int_{t'}^{t''} \alpha(t, x)\, dw(t)\right\}$$

uniformément en x sur tout intervalle fini, si seulement l_n est tel que $t_{nl_n} \leqslant t'$.

En effet, comme dans la démonstration du lemme 1, on trouve que

$$\mathsf{E}\left\{\exp\left(i\lambda \sum_{t' \leqslant t_{nk} \leqslant t''} \alpha(t_{nk}, x)\, \omega_{nk} \mid \mathfrak{F}_{l_n}\right)\right\} - \\ - \exp\left\{-\frac{\lambda^2}{2} \sum_{t' \leqslant t_{nk} \leqslant t''} \alpha^2(t_{nk}, x)\, \Delta t_{nk}\right\} \to 0$$

uniformément en x sur tout intervalle fini, car $\alpha(t, x)$ est bornée sur tout intervalle fini de variation de x.

Reste à remarquer que

$$\sum_{t' \leqslant t_{nk} \leqslant t''} \alpha^2(t_{nk}, x)\, \Delta t_{nk} \to \int_{t'}^{t''} \alpha^2(t, x)\, dt$$

uniformément en x sur tout intervalle fini, car $\alpha(t, x)$ est uniformément continue en t et x sur tout intervalle fini de variation de x.

LEMME 2. *Si une suite de fonctions mesurables $\varphi_n(x_1, \ldots, x_m)$ est bornée par une constante et converge uniformément vers une fonction continue $\varphi_0'(x_1, \ldots, x_m)$ sur tout compact, et si la suite de fonctions de répartition $F_n(x_1, \ldots, x_m)$ converge faiblement vers la fonction $F_0(x_1, \ldots, x_m)$, alors*

$$\lim_{n\to\infty} \int \varphi_n(x_1, \ldots, x_m)\, dF_n(x_1, \ldots, x_m) = \\ = \int \varphi_0(x_1, \ldots, x_m)\, dF_0(x_1, \ldots, x_m).$$

Démonstration. Il suffit de prouver que

$$\overline{\lim_{n\to\infty}} \int |\varphi_n(x_1, \ldots, x_m) - \varphi_0(x_1, \ldots, x_m)|\, dF_n(x_1, \ldots, x_m) = 0,$$

puisque

$$\lim_{n\to\infty} \int \varphi_0(x_1, \ldots, x_m)\, dF_n(x_1, \ldots, x_m) = \\ = \int \varphi_0(x_1, \ldots, x_m)\, dF_0(x_1, \ldots, x_m)$$

en vertu de la faible convergence de $F_n(x_1, \ldots, x_m)$ vers $F_0(x_1, \ldots, x_m)$.

Or $\forall K > 0$

$$\int |\varphi_0 - \varphi_n| \, dF_n \leqslant \int\limits_{\sum_i |x_i| \leqslant K} |\varphi_0 - \varphi_n| \, dF_n + \int\limits_{\sum_i |x_i| > K} |\varphi_0 - \varphi_n| \, dF_n ;$$

la première intégrale tend vers zéro pour $n \to \infty$, car $|\varphi_0 - \varphi_n|$ tend uniformément vers zéro pour $\sum |x_i| \leqslant K$; la deuxième intégrale peut être rendue aussi petite que l'on veut pour tous les n par le choix de K assez grand, car $|\varphi_0 - \varphi_n|$ est bornée et la suite de répartitions F_n faiblement convergente.

LEMME 3. *Soient* $\xi_1^{(n)}, \ldots, \xi_m^{(n)}$ $(n = 0, 1, \ldots)$ *m suites de variables aléatoires et* $\Phi_k^{(n)}(\lambda, x_1, \ldots, x_{k-1})$ *des fonctions telles que presque sûrement*

$$\Phi_1^{(n)}(\lambda) = \mathsf{E} e^{i\lambda\xi_1^{(n)}}, \quad \Phi_k^{(n)}(\lambda, \xi_1^{(n)}, \ldots, \xi_{k-1}^{(n)}) =$$

$$= \mathsf{E}\,(e^{i\lambda\xi_k^{(n)}} \mid \xi_1^{(n)}, \ldots, \xi_{k-1}^{(n)}), \quad k = 2, \ldots, m.$$

Si les fonctions $\Phi_k^{(0)}(\lambda, x_1, \ldots, x_{k-1})$ *sont continues pour tous les* k *et* $\Phi_k^{(n)}(\lambda, x_1, \ldots, x_{k-1}) \to \Phi_k^{(0)}(\lambda, x_1, \ldots, x_{k-1})$, $k = 1, 2, \ldots, m$, *pour* $n \to \infty$ *uniformément sur tout compact, alors la répartition conjointe de* $\xi_1^{(n)}, \ldots, \xi_m^{(n)}$ *converge faiblement vers la répartition conjointe de* $\xi_1^{(0)}, \ldots, \xi_m^{(0)}$.

Démonstration. Pour $k = 1$ on constate que la répartition de la variable $\xi_1^{(n)}$ converge vers la répartition de $\xi_1^{(0)}$. Supposons que la répartition conjointe $F_{k-1}^{(n)}$ des variables $\xi_1^{(n)}, \ldots, \xi_{k-1}^{(n)}$ tend vers la répartition conjointe $F_{k-1}^{(0)}$ des variables $\xi_1^{(0)}, \ldots, \xi_{k-1}^{(0)}$.

Le lemme 2 nous dit que

$$\mathsf{E} \exp\left\{i \sum_{j=1}^{k} \lambda_j \xi_j^{(n)}\right\} =$$

$$= \int \exp\left\{i \sum_{j=1}^{k-1} \lambda_j x_j\right\} \Phi_k^{(n)}(\lambda_k, x_1, \ldots, x_{k-1}) \, dF_{k-1}^{(n)}(x_1, \ldots, x_{k-1}) \to$$

$$\to \int \exp\left\{i \sum_{j=1}^{k-1} \lambda_j x_j\right\} \Phi_k^{(0)}(\lambda_k, x_1, \ldots, x_{k-1}) \, dF_{k-1}^{(0)}(x_1, \ldots, x_{k-1}) =$$

$$= \mathsf{E} \exp\{i\lambda_1\xi_1^{(0)} + i\lambda_2\xi_2^{(0)} + \ldots + i\lambda_k\xi_k^{(0)}\}.$$

Donc la répartition conjointe des variables $\xi_1^{(n)}, \ldots, \xi_k^{(n)}$ tendra également vers celle des variables $\xi_1^{(0)}, \ldots, \xi_k^{(0)}$. En raisonnant par récurrence sur k on obtient ce qu'on voulait. ■

Revenons à la démonstration du théorème. Soit une partition quelconque de l'intervalle $[0, 1]: 0 = \tau_0 < \tau_1 < \ldots < \tau_N = 1$. Posons

$$\eta_n^*(\tau_0) = \xi_{n0},$$

$$\eta_n^*(\tau_k) = \eta_n^*(\tau_{k-1}) + \sum_{\tau_{k-1} \leq t_{nr} < \tau_k} a(t_{nr}, \eta_n^*(\tau_{k-1}))\,\Delta t_{nr} + \sum_{\tau_{k-1} \leq t_{nr} < \tau_k} \sigma(t_{nr}, \eta_n^*(\tau_{k-1}))\,\omega_{nr}, \quad k = 1, \ldots, N,$$

$$\eta_0^*(\tau_0) = \xi_0,$$

$$\eta_0^*(\tau_k) = \eta_0^*(\tau_{k-1}) + \int_{\tau_{k-1}}^{\tau_k} a(s, \eta_0^*(\tau_{k-1}))\,ds + \int_{\tau_{k-1}}^{\tau_k} \sigma(s, \eta_0^*(\tau_{k-1}))\,dw(s), \quad k = 1, \ldots, N.$$

Il est évident que

$$\mathsf{E}\left(e^{i\lambda\eta_0^*(\tau_k)} \mid \eta_0^*(\tau_0), \ldots, \eta_0^*(\tau_{k-1})\right) = \Phi_k^{(0)}(\lambda, \eta_0^*(\tau_{k-1})),$$

où

$$\Phi_k^{(0)}(\lambda, x) = \exp\left\{i\lambda x + i\lambda \int_{\tau_{k-1}}^{\tau_k} a(s, x)\,ds - \frac{\lambda^2}{2} \int_{\tau_{k-1}}^{\tau_k} \sigma^2(s, x)\,ds\right\}$$

est une fonction continue de x. De la remarque du lemme 1 il résulte que

$$\mathsf{E}\left(e^{i\lambda\eta_n^*(\tau_k)} \mid \eta_n^*(\tau_0), \ldots, \eta_n^*(\tau_{k-1})\right) =$$
$$= \mathsf{E}\left\{\exp\left(i\lambda\eta_n^*(\tau_{k-1}) + i\lambda \sum_{\tau_{k-1} \leq t_{nr} < \tau_k} a(t_{nr}, \eta_n^*(\tau_{k-1}))\,\Delta t_{nr}\right) \times\right.$$
$$\left.\times\, \mathsf{E}\left[\exp\left(i\lambda \sum_{\tau_{k-1} \leq t_{nr} < \tau_n} \sigma(t_{nr}, \eta_n^*(\tau_{k-1}))\,\omega_{nr}\right) \mid \eta_n^*(\tau_0), \ldots, \eta_n^*(\tau_{k-1})\right]\right\}$$

tendra vers $\Phi_k^{(0)}(\lambda, x)$ uniformément en x sur tout intervalle fini si l'on substitue x à $\eta_n^*(\tau_{k-1})$. Donc en vertu du lemme 3 la répartition conjointe des variables $\eta_n^*(\tau_0), \ldots, \eta_n^*(\tau_N)$ convergera vers la répartition conjointe des variables $\eta_0^*(\tau_0,), \ldots, \eta_0^*(\tau_N)$.

Comme

$$\xi(\tau_k) - \eta_0^*(\tau_k) = \xi(\tau_{k-1}) - \eta_0^*(\tau_{k-1}) + \int_{\tau_{k-1}}^{\tau_k} [a(s, \xi(s)) - a(s, \eta_0^*(\tau_{k-1}))]\,ds +$$

$$+\int_{\tau_{k-1}}^{\tau_k} [\sigma(s, \xi(s)) - \sigma(s, \eta_0^*(\tau_{k-1}))]\, dw(s),$$

la condition de Lipschitz et quelques transformations simples donnent

$$\mathsf{E}\,|\,\xi(\tau_k) - \eta_0^*(\tau_k)\,|^2 = \mathsf{E}\,|\,\xi(\tau_{k-1}) - \eta_0^*(\tau_{k-1})\,|^2 +$$
$$+ 2\mathsf{E}\,|\,\xi(\tau_{k-1}) - \eta_0^*(\tau_{k-1})\,| \int_{\tau_{k-1}}^{\tau_k} |\, a(s, \xi(s)) - a(s, \eta_0^*(\tau_{k-1}))\,|\, ds +$$
$$+ \mathsf{E}\left(\int_{\tau_{k-1}}^{\tau_k} (a(s, \xi(s)) - a(s, \eta_0^*(\tau_{k-1})))\, ds\right)^2 +$$
$$+ \mathsf{E}\left(\int_{\tau_{k-1}}^{\tau_k} (\sigma(s, \xi(s)) - \sigma(s, \eta_0^*(\tau_{k-1})))\, dw(s)\right)^2 \leqslant$$
$$\leqslant \mathsf{E}\,|\,\xi(\tau_{k-1}) - \eta_0^*(\tau_{k-1})\,|^2 +$$
$$+ 2K\mathsf{E}\,|\,\xi(\tau_{k-1}) - \eta_0^*(\tau_{k-1})\,| \int_{\tau_{k-1}}^{\tau_k} |\,\xi(s) - \eta_0^*(\tau_{k-1})\,|\, ds +$$
$$+ 2\mathsf{E} \int_{\tau_{k-1}}^{\tau_k} ([a(s, \xi(s)) - a(s, \eta_0^*(\tau_{k-1}))]^2 +$$
$$+ [\sigma(s, \xi(s)) - \sigma(s, \eta_0^*(\tau_{k-1}))]^2)\, ds \leqslant$$
$$\leqslant \mathsf{E}\,|\,\xi(\tau_{k-1}) - \eta_0^*(\tau_{k-1})\,|^2 (1 + H(\tau_k - \tau_{k-1})) +$$
$$+ H \int_{\tau_{k-1}}^{\tau_k} |\,\xi(s) - \xi(\tau_{k-1})\,|^2\, ds,$$

où $H = 2K + 8K^2$. Donc, en utilisant les majorations précédentes, on trouve

$$\mathsf{E}\,(\xi(\tau_k) - \eta_0^*(\tau_k))^2 \leqslant He^H \mathsf{E} \sum_{k=1}^{N} \int_{\tau_{k-1}}^{\tau_k} |\,\xi(s) - \xi(\tau_{k-1})\,|^2\, ds.$$

Le corollaire 2 du § 2, chapitre VIII, nous apprend que

$$\mathsf{E}\,|\,\xi(s_1) - \xi(s_2)\,|^2 \leqslant C\,|\,s_1 - s_2\,|,$$

de sorte que

$$\mathsf{E}\,(\xi(\tau_k)-\eta_0^*(\tau_k))^2=O\,[\max_k(\tau_k-\tau_{k-1})]. \qquad (3)$$

Par des raisonnements analogues on établit que

$$\overline{\lim_{n\to\infty}}\,\mathsf{E}\,[\eta_n(\tau_k)-\eta_n^*(\tau_k)]^2=O\,(\max_k(\tau_k-\tau_{k-1})). \qquad (4)$$

A noter que quels que soient les nombres réels $\lambda_1, \ldots, \lambda_m$ et les variables aléatoires $\xi_1, \ldots, \xi_m, \eta_1, \ldots, \eta_m$, on a

$$|\,\mathsf{E}\exp(i\lambda_1\xi_1+\ldots+i\lambda_m\xi_m)-\mathsf{E}\exp(i\lambda_1\eta_1+\ldots+i\lambda_m\eta_m)\,|\leqslant$$
$$\leqslant\mathsf{E}\,|\exp\{i\lambda_1(\xi_1-\eta_1)+\ldots+i\lambda_m(\xi_m-\eta_m)\}-1\,|\leqslant$$
$$\leqslant\mathsf{E}\,|\,\lambda_1(\xi_1-\eta_1)+\ldots+\lambda_m(\xi_m-\eta_m)\,|\leqslant$$
$$\leqslant\sqrt{\sum_{k=1}^m\lambda_k^2\sum_{k=1}^m\mathsf{E}\,(\xi_k-\eta_k)^2}. \qquad (5)$$

Soient $s_1, \ldots, s_m$ des nombres arbitraires de $[0, 1]$. On peut admettre que $s_i=\tau_{k_i}$. Alors

$$\overline{\lim_{n\to\infty}}\,|\,\mathsf{E}\exp\{i\sum\lambda_j\eta_n(\tau_{k_j})\}-\mathsf{E}\exp\{i\sum\lambda_j\xi(\tau_{k_j})\}\,|\leqslant$$
$$\leqslant\overline{\lim_{n\to\infty}}\,|\,\mathsf{E}\exp\{i\sum\lambda_j\eta_n(\tau_{k_j})\}-\mathsf{E}\exp\{i\sum\lambda_j\eta_n^*(\tau_{k_j})\}\,|+$$
$$+\overline{\lim_{n\to\infty}}\,|\,\mathsf{E}\exp\{i\sum\lambda_j\eta_n^*(\tau_{k_j})\}-\mathsf{E}\exp\{i\sum\lambda_j\eta_0^*(\tau_{k_j})\}\,|+$$
$$+\,|\,\mathsf{E}\exp\{i\sum\lambda_j\eta_0^*(\tau_{k_j})\}-\mathsf{E}\exp\{i\sum\lambda_j\xi(\tau_{k_j})\}\,|\leqslant$$
$$\leqslant O\left(\sqrt{m\max_k(\tau_k-\tau_{k-1})}\right)$$

en vertu des relations (3), (4), (5) et de la convergence de la répartition conjointe des variables $\eta_n^*(s_j)$ vers celle des $\eta_0^*(s_j)$. Ce qui achève la démonstration du théorème, car $\max_k(\tau_k-\tau_{k-1})$ arbitraire. ■

THÉORÈME 2. *Dans les hypothèses du théorème* 1, *la répartition de* $f(\xi_n(t))$ *convergera vers celle de* $f(\xi(t))$, $\forall f\in\mathscr{C}_{[0,1]}$.

D é m o n s t r a t i o n. La convergence des répartitions finidimensionnelles des processus $\xi_n(t)$ vers celles de $\xi(t)$ et la remarque du § 2 nous permettent de ramener la démonstration du théorème à celle de la relation

$$\lim_{h\to\infty}\overline{\lim_{n\to\infty}}\,\mathsf{P}\{\sup_{|t'-t''|\leqslant h}|\,\xi_n(t')-\xi_n(t'')\,|>\varepsilon\}=0. \qquad (6)$$

Les raisonnements du théorème 2, § 2, donnent

$$\mathsf{P}\{\sup_{|t'-t''|\leqslant h} |\xi_n(t')-\xi_n(t'')|>\varepsilon\}\leqslant$$
$$\leqslant \sum_{kh<1} \mathsf{P}\left\{\sup_{kh\leqslant t\leqslant(k+1)h} |\xi_n(t)-\xi_n(kh)|>\frac{\varepsilon}{4}\right\}.$$

Désignons par s_k le plus grand des indices r tels que $t_{nr}\leqslant kh$. On a alors

$$\mathsf{P}\left\{\sup_{kh\leqslant t\leqslant(k+1)h} |\xi_n(t)-\xi_n(kh)|>\frac{\varepsilon}{4}\right\}\leqslant$$
$$\leqslant \mathsf{P}\left\{\sup_{s_k\leqslant j_1<j_2\leqslant s_{k+1}+1} |\xi_{nj_1}-\xi_{nj_2}|>\frac{\varepsilon}{4}\right\}\leqslant$$
$$\leqslant \mathsf{P}\left\{\sup_{s_k\leqslant j\leqslant s_{k+1}+1} |\xi_{nj}-\xi_{ns_k}|>\frac{\varepsilon}{8}\right\}.$$

Pour estimer cette probabilité nous aurons besoin du

LEMME 4. *Si $\xi_1, \xi_2, \ldots, \xi_m$ forment une chaîne de Markov et*

$$\mathsf{P}\{|\xi_m-\xi_k|>c\mid\xi_k\}\leqslant\alpha, \text{ où } \alpha<1,$$

presque sûrement, alors

$$\mathsf{P}\{\sup_k |\xi_k|>2c\}\leqslant\frac{1}{1-\alpha}\mathsf{P}\{|\xi_m|>c\}.$$

D é m o n s t r a t i o n.

$$\mathsf{P}\{\sup_k |\xi_k|>2c,\ |\xi_m|\leqslant c\}=$$
$$=\sum_{k=1}^{m}\mathsf{P}\{|\xi_i|\leqslant 2c,\ i\leqslant k-1,\ |\xi_k|>2c,\ |\xi_m|\leqslant c\}\leqslant$$
$$\leqslant\sum_{k=1}^{m}\mathsf{P}\{|\xi_i|\leqslant 2c,\ i\leqslant k-1,\ |\xi_k|>2c,\ |\xi_m-\xi_k|>c\}=$$
$$=\sum_{k=1}^{m}\mathsf{E}(\mathsf{P}\{|\xi_i|\leqslant 2c,\ i\leqslant k-1,\ |\xi_k|>2c\mid\xi_k\}\mathsf{P}\{\xi_m-\xi_k|>c\mid\xi_k\})\leqslant$$
$$\leqslant\alpha\sum_{k=1}^{m}\mathsf{E}\mathsf{P}\{|\xi_i|\leqslant 2c,\ i\leqslant k-1,\ |\xi_k|>2c\mid\xi_k\}=$$
$$=\alpha\mathsf{P}\{\sup_k |\xi_k|>2c\}.$$

Donc

$$\mathsf{P}\{\sup_k |\xi_k|>2c\}\leqslant\mathsf{P}\{|\xi_m|>c\}+\mathsf{P}\{\sup_k |\xi_k|>2c,\ |\xi_m|\leqslant c\}\leqslant$$
$$\leqslant\mathsf{P}\{|\xi_m|>c\}+\alpha\mathsf{P}\{\sup_k |\xi_k|>2c\}.$$

Le lemme est une conséquence de la dernière relation. ■

Revenons à la démonstration du théorème. Comme

$$\mathsf{P}\{|\xi_{ns_{k+1}+1}-\xi_{nj}|>\delta\mid\xi_{nj}\}\leqslant\frac{1}{\delta^2}\mathsf{E}\left([\xi_{ns_{k+1}+1}-\xi_{nj}]^2\mid\xi_{nj}\right)=$$

$$=\frac{1}{\delta^2}\mathsf{E}\left\{\left(\sum_{r=j}^{s_{k+1}}[a_n(t_{nr},\xi_{nr})\Delta t_{nr}+\sigma_n(t_{nr},\xi_{nr})\omega_{nr}]\right)^2\Big|\xi_{nj}\right\}\leqslant$$

$$\leqslant\frac{2}{\delta^2}\mathsf{E}\left[\left(\sum_{r=j}^{s_{k+1}}a_n(t_{nr},\xi_{nr})\Delta t_{nr}\right)^2\Big|\xi_{nj}\right]+$$

$$+\frac{2}{\delta^2}\mathsf{E}\left[\left(\sum_{r=j}^{s_{k+1}}\sigma_n(t_{nr},\xi_{nr})\omega_{nr}\right)^2\Big|\xi_{nj}\right]\leqslant$$

$$\leqslant\frac{2}{\delta^2}\mathsf{E}\left[\left(\sum_{r=j}^{s_{k+1}}a_n(t_{nr},\xi_{nr})\Delta t_{nr}\right)^2\Big|\xi_{nj}\right]+$$

$$+\frac{2}{\delta^2}\mathsf{E}\left[\sum_{r=j}^{s_{k+1}}\sigma_n^2(t_{nr},\xi_{nr})\mathsf{E}(\omega_{nr}^2\mid\xi_{nr})\mid\xi_{nj}\right]+$$

$$+\frac{4}{\delta^2}\mathsf{E}\left(\sum_{j=r<l<s_{k+1}}\sigma_n(t_{nr},\xi_{nr})\sigma_n(t_{nl},\xi_{nl})\omega_{nr}\mathsf{E}(\omega_{nl}\mid\xi_{nl})\mid\xi_{nj}\right)=$$

$$=\mathsf{E}\left[\left(\sum_{r=j}^{s_{k+1}}a_n(t_{nr},\xi_{nr})\Delta t_{nr}\right)^2+\sum_{r=j}^{s_{k+1}}\sigma_n^2(t_{nr},\xi_{nr})\Delta t_{nr}\mid\xi_{nj}\right],$$

que les fonctions a_n et σ_n sont bornées et enfin que

$$\sum_{r=s_k}^{s_{k+1}}\Delta t_{nj}\leqslant h+2\max_j\Delta t_{nj},$$

il existe une constante H_1 telle que

$$\mathsf{P}\{|\xi_{ns_{k+1}+1}-\xi_{kj}|>\delta\mid\xi_{nj}\}\leqslant\frac{H_1}{\delta^2}(h+2\max_j\Delta t_{nj}).$$

Donc, pour h assez petit et tous les n assez grands

$$\mathsf{P}\{|\xi_{ns_{k+1}+1}-\xi_{nj}|>\varepsilon/16\}\leqslant 1/2$$

et en vertu du lemme 4

$$\mathsf{P}\{\sup_{kh\leqslant t\leqslant(k+1)h}|\xi_n(t)-\xi_n(kh)|>\varepsilon/4\}\leqslant 2\mathsf{P}\{|\xi_{ns_{k+1}+1}-\xi_{s_k}|>\varepsilon/16\}.$$

La convergence des répartitions de $\xi_n(t)$ vers celles de $\xi(t)$ implique

$$\varlimsup_{n\to\infty} \mathsf{P}\{|\xi_{ns_{k+1}+1}-\xi_{ns_k}|>\varepsilon/16\}\leqslant \mathsf{P}\{|\xi((k+1)h)-\xi(kh)|>\varepsilon/16\}.$$

Par suite

$$\varlimsup_{n\to\infty} \mathsf{P}\{\sup_{|t'-t''|\leqslant h} |\xi_n(t')-\xi_n(t'')|>\varepsilon\}\leqslant$$
$$\leqslant \sum_{kh<1} \mathsf{P}\{|\xi(kh+h)-\xi(kh)|>\varepsilon/16\}\leqslant$$
$$\leqslant \sum_{kh<1} \left(\frac{16}{\varepsilon}\right)^4 \mathsf{E}\,|\xi(kh+h)-\xi(kh)|^4.$$

Or, d'après le corollaire 2, § 2, chapitre VIII,

$$\mathsf{E}\,|\xi(t+h)-\xi(t)|^4\leqslant Lh^2$$

pour un L, donc

$$\varlimsup_{n\to\infty} \mathsf{P}\{\sup_{|t'-t''|\leqslant h} |\xi_n(t')-\xi_n(t'')|>\varepsilon\}\leqslant \left(\frac{16}{\varepsilon}\right)^4 Lh.$$

Ce qui démontre (6) et partant le théorème. ■

§ 5. L'espace des fonctions ne présentant pas de discontinuités de seconde espèce

Appelons $\mathscr{D}_{[0,1]}$ l'ensemble des fonctions $x(t)$, définies sur l'intervalle [0, 1], à valeurs réelles et admettant des dérivées à gauche et à droite en chaque point. On identifiera les fonctions qui sont confondues en tous les points de continuité; il faudra donc convenir d'une définition standard des valeurs de $x(t)$ aux points de discontinuité. Dans la suite on admettra que toutes les fonctions de $\mathscr{D}_{[0,1]}$ sont telles que

$$x(t)=x(t+0),\ x(0)=x(+0),\ x(1)=x(1-0). \qquad (1)$$

L'étude de l'espace $\mathscr{D}_{[0,1]}$ est intéressante, car il existe des classes de processus aléatoires dont les réalisations ne présentent presque sûrement pas de discontinuités de seconde espèce (par exemple, les processus à accroissements indépendants, les processus markoviens sous des hypothèses assez larges). Pour pouvoir se servir des résultats du § 1, il faut munir $\mathscr{D}_{[0,1]}$ d'une métrique qui le transforme en un espace métrique séparable dont la plus petite tribu des ensembles cylindriques soit confondue avec la tribu de ses boréliens. Cette métrique doit être assez « forte » (de façon à ce qu'il y ait le moins possible de suites convergentes, et partant le plus possible de fonctionnelles continues dans cette métrique). La métrique uniforme

$$\rho_u(x, y)=\sup_{0\leqslant t\leqslant 1} |x(t)-y(t)|$$

ne convient pas, puisque l'espace $\mathscr{D}_{[0,1]}$ ne sera pas séparable (l'ensemble des fonctions $x_s(t) = \frac{1+\operatorname{sgn}(t-s)}{2}$, $0 < s < 1$, possède la puissance du continu, mais la distance de deux quelconques des éléments de cet ensemble est égale à 1). Munissons l'espace $\mathscr{D}_{[0,1]}$ d'une métrique plus faible que la métrique uniforme.

Appelons Λ l'ensemble de toutes les fonctions numériques $\lambda(t)$ monotones croissantes sur $[0, 1]$, telles que $\lambda(0) = 0$, $\lambda(1) = 1$ (i.e. λ définit une bijection de $[0, 1]$ sur lui-même).

A noter que $\lambda^{-1} \in \Lambda$, $\forall \lambda \in \Lambda$. Si λ_1 et $\lambda_2 \in \Lambda$, alors la fonction composée $\lambda_1(\lambda_2(\cdot)) \in \Lambda$.

Définissons maintenant pour tout couple $x(t)$ et $y(t)$ de $\mathscr{D}_{[0,1]}$ la quantité

$$\rho_{\mathscr{D}}(x, y) = \inf_{\lambda \in \Lambda} [\sup_t |x(t) - y(\lambda(t))| + \sup_t |t - \lambda(t)|]. \tag{2}$$

Nous allons montrer que $\rho_{\mathscr{D}}$ définit une métrique dans $\mathscr{D}_{[0,1]}$ Pour cela il nous faut vérifier que la fonction $\rho_{\mathscr{D}}$ satisfait les trois axiomes suivants:

a) $\rho_{\mathscr{D}}(x, y) \geqslant 0$, l'égalité n'étant réalisée que pour $x = y$;

b) $\rho_{\mathscr{D}}(x, y) = \rho_{\mathscr{D}}(y, x)$;

c) $\rho_{\mathscr{D}}(x, z) \leqslant \rho_{\mathscr{D}}(x, y) + \rho_{\mathscr{D}}(y, z)$ pour tous les $x(t)$, $y(t)$ et $z(t)$ de $\mathscr{D}_{[0,1]}$.

La propriété a) est évidente. La propriété b) découle de la relation

$$\rho_{\mathscr{D}}(y, x) = \inf_{\lambda \in \Lambda} [\sup_t |y(t) - x(\lambda(t))| + \sup_t |t - \lambda(t)|] =$$
$$= \inf_{\lambda^{-1} \in \Lambda} [\sup_t |y(\lambda^{-1}(t)) - x(t)| + \sup_t |\lambda^{-1}(t) - t|] = \rho_{\mathscr{D}}(x, y).$$

Attardons-nous sur la propriété c) dite encore inégalité du triangle. Soient $x(t)$, $y(t)$ et $z(t)$ des fonctions de $\mathscr{D}_{[0,1]}$. Pour tout $\varepsilon > 0$ il existe des fonctions $\lambda_1(t)$ et $\lambda_2(t)$ telles que

$$\begin{aligned} \rho_{\mathscr{D}}(x, y) &\geqslant \sup_t |x(t) - y(\lambda_1(t))| + \sup_t |t - \lambda_1(t)| - \varepsilon, \\ \rho_{\mathscr{D}}(y, z) &\geqslant \sup_t |y(t) - z(\lambda_2(t))| + \sup_t |t - \lambda_2(t)| - \varepsilon. \end{aligned} \tag{3}$$

On a

$$\rho_{\mathscr{D}}(x, z) \leqslant \sup_t |x(t) - z(\lambda_2(\lambda_1(t)))| + \sup_t |t - \lambda_2(\lambda_1(t))| \leqslant$$
$$\leqslant \sup_t |x(t) - y(\lambda_1(t))| + \sup_t |t - \lambda_1(t)| +$$
$$+ \sup_t |y(\lambda_1(t)) - z(\lambda_2(\lambda_1(t)))| + \sup_t |\lambda_1(t) - \lambda_2(\lambda_1(t))| =$$

$$= \sup_t |x(t) - y(\lambda_1(t))| + \sup_t |t - \lambda_1(t)| +$$
$$+ \sup_t |y(t) - z(\lambda_2(t))| + \sup_t |t - \lambda_2(t)|,$$

puisque si t parcourt $[0, 1]$, alors $\lambda_1(t)$ parcourra également l'intervalle $[0, 1]$. Compte tenu de (3) on obtient

$$\rho_{\mathscr{D}}(x, z) \leqslant \rho_{\mathscr{D}}(x, y) + \rho_{\mathscr{D}}(y, z) + 2\varepsilon,$$

d'où la propriété c), car $\varepsilon > 0$ arbitraire.

Posons

$$\Delta_c(x) = \sup_{t-c\leqslant t'\leqslant t\leqslant t''\leqslant t+c} [\min\{|x(t') - x(t)|\,;\ |x(t'') - x(t)|\}] +$$
$$+ \sup_{0\leqslant t\leqslant c} |x(t) - x(0)| + \sup_{1-c\leqslant t\leqslant 1} |x(t) - x(1)|.$$

Il est immédiat que

$$\lim_{c\to 0} \Delta_c(x) = 0$$

pour $x(t) \in \mathscr{D}_{[0,1]}$.

Lemme 1. *Soit* $x(t) \in \mathscr{D}_{[0,1]}$ *et* $[\alpha, \beta] \subset [0, 1]$. *Si* $x(t)$ *ne présente pas de sauts* $> \varepsilon$ *sur* $[\alpha, \beta]$, *alors*

$$|x(t') - x(t'')| \leqslant 2\Delta_c(x) + \varepsilon$$

pour $|t' - t''| \leqslant c$, $t', t'' \in [\alpha, \beta]$.

D é m o n s t r a t i o n. Prenons un $\delta \in]0, \varepsilon[$ et un point $\tau \in [t', t'']$ tel que

$$\text{pour} \quad t \in [t', \tau[\begin{cases} |x(t') - x(t)| < \Delta_c(x) + \delta, \\ |x(t') - x(\tau)| \geqslant \Delta_c(x) + \delta. \end{cases}$$

Si un tel point n'existe pas, alors $|x(t') - x(t'')| \leqslant \Delta_c(x) + \delta$ et le lemme est vrai. Si τ existe, les inégalités

$$\min[|x(\tau) - x(t')|;\ |x(\tau) - x(t'')|] \leqslant \Delta_c(x),$$
$$|x(\tau) - x(t')| \geqslant \Delta_c(x) + \delta$$

entraînent

$$|x(\tau) - x(t'')| \leqslant \Delta_c(x).$$

Donc

$$|x(t'') - x(t')| \leqslant |x(t'') - x(\tau)| + |x(\tau) - x(\tau - 0)| +$$
$$+ |x(\tau - 0) - x(t')| \leqslant 2\Delta_c(x) + \delta + \varepsilon.$$

D'où le lemme en passant à la limite pour $\delta \downarrow 0$. ■

Appelons $H_{m,n}$ l'ensemble des fonctions $x(t)$ de $\mathscr{D}_{[0,1]}$ constantes sur tout intervalle $\left[\frac{k}{n}, \frac{k+1}{n}\right[$ et prenant des valeurs multiples de m.

LEMME 2. *Pour toute fonction* $x(t) \in \mathscr{D}_{[0,1]}$ *il existe une fonction* $x^*(t) \in H_{m,n}$ *telle que*

$$\rho_{\mathscr{D}}(x, x^*) \leqslant \frac{1}{n} + \frac{1}{m} + 4\Delta_{\frac{2}{n}}(x). \qquad (4)$$

Démonstration. Dans chaque intervalle $\left]\frac{k}{n}, \frac{k+1}{n}\right[$ il existe au plus un point en lequel le saut est $> 2\Delta_{\frac{2}{n}}(x)$. En effet, si τ est un tel point, alors

$$|x(s) - x(\tau - 0)| = \min[|x(s) - x(\tau - 0)|;$$
$$|x(\tau) - x(\tau - 0)|] \leqslant \Delta_{\frac{1}{n}}(x) \quad \text{pour} \quad s \in \left[\frac{k}{n}, \tau\right[,$$
$$|x(s) - x(\tau)| \leqslant \Delta_{\frac{1}{n}}(x) \quad \text{pour} \quad s \in \left[\tau, \frac{k+1}{n}\right]$$

et par suite

$$|x(s) - x(s-0)| \leqslant 2\Delta_{\frac{1}{n}}(x) \leqslant 2\Delta_{\frac{2}{n}}(x).$$

Soit τ_k le point (s'il existe) de $\left[\frac{k}{n}, \frac{k+1}{n}\right]$ en lequel

$$|x(\tau_k) - x(\tau_k - 0)| \geqslant 2\Delta_{\frac{2}{n}}(x).$$

Soit $\lambda(t) \in \Lambda$ telle que

$$\lambda\left(\frac{k+1}{n}\right) = \tau_k \quad \text{et} \quad t - \frac{1}{n} \leqslant \lambda(t) \leqslant t$$

(par exemple, la fonction linéaire par morceaux, définie par les égalités $\lambda(0) = 0$, $\lambda\left(\frac{k+1}{n}\right) = \tau_k$, $\lambda(1) = 1$). Posons $\bar{x}(t) = x(\lambda(t))$. La fonction $\bar{x}(t)$ présentera des sauts $> 2\Delta_{\frac{2}{n}}(x)$ uniquement aux points de la forme $\frac{k}{n}$ et

$$\rho_{\mathscr{D}}(x, \bar{x}) \leqslant \sup_t |\bar{x}(t) - x(\lambda(t))| + \sup_t |t - \lambda(t)| \leqslant \frac{1}{n}$$

Soit d'autre part

$$\bar{x}^*(t) = \begin{cases} \bar{x}\left(\frac{k}{n}\right) & \text{pour } t \in \left[\frac{k}{n}, \frac{k+1}{n}\right[, \quad k \leqslant n-1; \\ \bar{x}\left(\frac{n-1}{n}\right) & \text{pour } t = 1. \end{cases}$$

Il vient

$$\rho_{\mathcal{D}}(\overline{x}, \overline{x}^*) \leqslant \sup |\overline{x}(t) - \overline{x}^*(t)| \leqslant \sup_k \sup_{\frac{k}{n} \leqslant t < \frac{k+1}{n}} \left| \overline{x}(t) - \overline{x}\left(\frac{k}{n}\right) \right|.$$

Les sauts de $x(t)$ supérieurs à $2\Delta_{\frac{2}{n}}(x)$ n'ayant lieu qu'aux points de la forme $\frac{k}{n}$, l'intervalle $\left[\frac{k}{n}, \frac{k+1}{n}\right[$ n'en contiendra pas et par suite le lemme 1 donne

$$\left| \overline{x}\left(\frac{k}{n}\right) - \overline{x}(t) \right| \leqslant 2\Delta_{\frac{1}{n}}(\overline{x}) + 2\Delta_{\frac{2}{n}}(x) \text{ pour } t \in \left[\frac{k}{n}, \frac{k+1}{n}\right[.$$

Evaluons $\Delta_{\frac{1}{n}}(\overline{x})$:

$$\Delta_{\frac{1}{n}}(\overline{x}) = \sup_{t-\frac{1}{n} \leqslant t' \leqslant t \leqslant t'' + \frac{1}{n}} [\min\{|\overline{x}(t') - \overline{x}(t)|\,;\ |\overline{x}(t) - \overline{x}(t'')|\}] +$$

$$+ \sup_{0 \leqslant t \leqslant \frac{1}{n}} |\overline{x}(t) - \overline{x}(0)| + \sup_{1-\frac{1}{n} \leqslant t \leqslant 1} |\overline{x}(t) - \overline{x}(1)| =$$

$$= \sup_{t-\frac{1}{n} \leqslant t' \leqslant t \leqslant t'' \leqslant t + \frac{1}{n}} [\min\{|x(\lambda(t')) - x(\lambda(t))|;\ |x(\lambda(t)) - x(\lambda(t''))|\}] +$$

$$+ \sup_{0 \leqslant t \leqslant \frac{1}{n}} |x(\lambda(t)) - x(0)| + \sup_{1-\frac{1}{n} \leqslant t \leqslant 1} |x(\lambda(t)) - x(1)|.$$

On remarquera que pour $t_1 < t_2 \leqslant t_1 + \frac{1}{n}$

$$t_1 - \frac{1}{n} \leqslant \lambda(t_1) < \lambda(t_2) \leqslant t_2 \leqslant t_1 + \frac{1}{n},$$

de sorte que $0 < \lambda(t_2) - \lambda(t_1) \leqslant \frac{2}{n}$. Donc $\Delta_{\frac{1}{n}}(\overline{x}) \leqslant \Delta_{\frac{2}{n}}(x)$ et par suite

$$\rho_{\mathcal{D}}(\overline{x}, \overline{x}^*) \leqslant 4\Delta_{\frac{2}{n}}(x).$$

Posons enfin $x^*(t) = \frac{1}{m}\operatorname{Ent}(m\overline{x}^*(t))$, où $\operatorname{Ent}(x)$ est la partie entière de x. L'inégalité

$$|x^*(t) - \overline{x}^*(t)| \leqslant \frac{1}{m}$$

entraîne

$$\rho_{\mathscr{D}}(x, x^*) \leqslant \rho_{\mathscr{D}}(x, \bar{x}) + \rho_{\mathscr{D}}(\bar{x}, \bar{x}^*) + \rho_{\mathscr{D}}(\bar{x}^*, x^*) \leqslant$$

$$\leqslant \frac{1}{n} + 4\Delta_{\frac{2}{n}}(x) + \frac{1}{m}. \blacksquare$$

COROLLAIRE. *L'espace* $\mathscr{D}_{[0,1]}$ *muni de la métrique* $\rho_{\mathscr{D}}$ *est séparable.*

THÉORÈME 1. *Soient L une constante positive,* $\omega(\delta)$ *une fonction définie, continue, monotone pour* $\delta > 0$ *et telle que* $\lim_{\delta \downarrow 0} \omega(\delta) = 0$. *Appelons* $K(L, \omega)$ *l'ensemble des fonctions de* $\mathscr{D}_{[0,1]}$ *telles que* $|x(t)| \leqslant L$, $\Delta_c(x) \leqslant \omega(c)$. *Alors* $K(L, \omega)$ *est un compact muni de la métrique* $\rho_{\mathscr{D}}$.

D é m o n s t r a t i o n. Remarquons tout d'abord que $K(L, \omega)$ possède un ε-réseau fini pour tout $\varepsilon > 0$; un tel ε-réseau est constitué par les fonctions de $H_{m,n}$ telles que $|x(t)| \leqslant L$ si m et n sont choisis tels que $\frac{1}{n} + \frac{1}{m} + 4\omega\left(\frac{2}{n}\right) < \varepsilon$. L'ensemble $K(L, \omega)$ est fermé. Il est aisé de vérifier que

$$\Delta_c(x) \leqslant \Delta_{c+\rho_{\mathscr{D}}(x, y)}(y) + 3\rho_{\mathscr{D}}(x, y).$$

Donc, si $\rho_{\mathscr{D}}(x_n, \bar{x}) \to 0$ et $x_n \in K(L, \omega)$, alors pour tout $\alpha > 0$

$$\Delta_c(\bar{x}) \leqslant \overline{\lim_{n \to \infty}} \Delta_{c+\alpha}(x_n) + 3\alpha \leqslant \omega(c + \alpha) + 3\alpha.$$

La continuité de ω implique $\Delta_c(\bar{x}) \leqslant \omega(c)$. Il est évident aussi

$$\sup_t |x(t)| \leqslant \sup_t |y(t)| + 3\rho_{\mathscr{D}}(x, y)$$

si bien que

$$\sup_t |\bar{x}(t)| \leqslant \overline{\lim_{n \to \infty}} [\sup_t |x_n(t)| + 3\rho_{\mathscr{D}}(x_n, \bar{x})] \leqslant L.$$

Donc la limite des suites de $K(L, \omega)$ appartiendra aussi à $K(L, \omega)$. Reste à prouver que toute suite de Cauchy $x_n(t) \in K(L, \omega)$ sera convergente (nous montrerons alors que $K(L, \omega)$ est un espace métrique complet possédant un ε-réseau pour tout $\varepsilon > 0$, or cela signifiera que $K(L, \omega)$ est un compact). Soit $x_n(t)$ suite de fonctions de $K(L, \omega)$ telle que $\rho_{\mathscr{D}}(x_n, x_m) \to 0$ pour n et $m \to \infty$ (i.e. $x_n(t)$ est une suite de Cauchy). Il suffit de prouver qu'une sous-suite $x_{n_k}(t)$ tend vers $\bar{x}(t)$. On peut donc admettre que la suite $x_n(t)$ est telle que $\rho_{\mathscr{D}}(x_n, x_{n+1}) < \frac{1}{2^{n+1}}$. Il existe alors une suite

de fonctions $\lambda_n(t) \in \Lambda$ telle que

$$\sup_t |x_n(t) - x_{n+1}(\lambda_{n+1}(t))| \leqslant \frac{1}{2^{n+1}}.$$

$$\sup_t |t - \lambda_{n+1}(t)| \leqslant \frac{1}{2^{n+1}}$$

Posons

$$\mu_1(t) = \lambda_1(t), \quad \mu_n(t) = \lambda_n(\mu_{n-1}(t)).$$

Comme

$$\sup_t |\mu_n(t) - \mu_{n-1}(t)| \leqslant \frac{1}{2^n},$$

$\mu_n(t)$ tend vers une fonction $\mu(t)$ continue croissante telle que $\mu(0) = 0$, $\mu(1) = 1$. Par ailleurs

$$\sup_t |x_n(\mu_n(t)) - x_{n-1}(\mu_{n-1}(t))| = |\sup_t |x_n(\lambda_n(t)) - x_{n-1}(t)| \leqslant \frac{1}{2^n}.$$

Donc $x_n(\mu_n(t))$ est uniformément convergente vers une fonction $x^*(t) \in \mathscr{D}_{[0,1]}$. Etudions le lien existant entre les fonctions $x^*(t)$ et $\mu(t)$. Soit $\mu(t)$ constante sur un intervalle $[\alpha, \beta]$.

Si $x^*(\alpha) = x^*(\beta)$, alors $x^*(t)$ est constante sur $[\alpha, \beta]$.

Si $x^*(\alpha) \neq x^*(\beta)$, il existe $\gamma \in [\alpha, \beta]$ tel que

$$x^*(t) = \begin{cases} x^*(\alpha) & \text{pour} \quad t \in [\alpha, \gamma[, \\ x^*(\beta) & \text{pour} \quad t \in [\gamma, \beta]. \end{cases}$$

En effet, le cas échéant il existerait des points $t' < t'' < t'''$ de $[\alpha, \beta]$ tels que $x^*(t') \neq x^*(t'')$, $x^*(t'') \neq x^*(t''')$ et

$$\lim_{n\to\infty} \min[|x_n(\mu_n(t')) - x_n(\mu_n(t''))|;\ |x_n(\mu_n(t'')) - x_n(\mu_n(t'''))|] =$$
$$= \min[|x^*(t') - x^*(t'')|;\ |x^*(t'') - x^*(t''')|] > 0,$$

bien que $\mu_n(t') < \mu_n(t'') < \mu_n(t''')$ et $\mu_n(t')$, $\mu_n(t'')$, $\mu_n(t''')$ tendent vers $\mu(\alpha)$. Ce qui contredirait l'appartenance de $x_n(t)$ à $K(L, \omega)$. Soit $\bar{x}(t) \in \mathscr{D}_{[0,1]}$ définie par la relation

$$x^*(t) = \bar{x}(\mu(t)) \tag{5}$$

qui est réalisée en tous les points t où $\mu(s) > \mu(t)$, $\forall s \in]t, 1]$. Il n'existe qu'une seule fonction $\bar{x}(t) \in \mathscr{D}_{[0,1]}$ vérifiant (5). Montrons que $\bar{x}(t)$ est la limite de la suite $x_n(t)$. Construisons à cet effet des fonctions auxiliaires $\varphi_n(t) \in \Lambda$. Soient $\tau_1, \tau_2, \ldots, \tau_k$ les points de $[0, 1]$ en lesquels $\bar{x}(t)$ présente des sauts supérieurs à $1/n$. Désignons par $[\alpha_i, \beta_i]$ le plus grand intervalle dans lequel $\mu(t)$ prend la valeur τ_i (cet intervalle peut être composé d'un seul point).

Soit $\gamma_i \in [\alpha_i, \beta_i]$ tel que

$$x^*(t) = \begin{cases} \bar{x}(\tau_i - 0) & \text{pour} \quad t \in [\alpha_i, \gamma_i[, \\ \bar{x}(\tau_i) & \text{pour} \quad t \in [\gamma_i, \beta_i] \end{cases}$$

(en particulier $x^*(t) = \bar{x}(\tau_i)$ sur $[\alpha_i, \beta_i]$ si $\alpha_i = \gamma_i$). Choisissons $\varepsilon_n \leqslant 1/n$ tel que $\Delta_{\varepsilon_n}(\bar{x}) < 1/n$. Soit $\varphi_n(t)$ une fonction telle que $\varphi_n(\gamma_i) = \tau_i$, $|\varphi_n(t) - \mu(t)| \leqslant \varepsilon_n$. Evaluons

$$\sup_t |x^*(t) - \bar{x}(\varphi_n(t))|.$$

Si $t \notin [\alpha_i, \beta_i]$, alors le lemme 1 nous dit que

$$|x^*(t) - \bar{x}(\varphi_n(t))| = |\bar{x}(\mu(t)) - \bar{x}(\varphi_n(t)) \leqslant 2\Delta_{\varepsilon_n}(\bar{x}) + 1/n,$$

car $\bar{x}(t)$ ne possède pas de sauts $> 1/n$ entre $\mu(t)$ et $\varphi_n(t)$.

Si $t \in [\alpha_i, \gamma_i[$, alors

$$|x^*(t) - \bar{x}(\varphi_n(t))| \leqslant \sup_{s \in [\tau_i - \varepsilon_n, \tau_i[} |\bar{x}(\tau_i - 0) - \bar{x}(s)| \leqslant \Delta_{\varepsilon_n}(\bar{x}),$$

puisque $|\bar{x}(\tau_i - 0) - \bar{x}(\tau_i)| > 1/n$.

Si $t \in [\gamma_i, \beta_i]$, on établit de façon analogue que

$$|x^*(t) - \bar{x}(\varphi_n(t))| \leqslant \Delta_{\varepsilon_n}(\bar{x}).$$

Donc

$$\sup_t |x^*(t) - \bar{x}(\varphi_n(t))| \leqslant 2\Delta_{\varepsilon_n}(\bar{x}) + 1/n \leqslant 3/n.$$

Evaluons maintenant $\rho_{\mathscr{D}}(x_n, \bar{x})$. On a

$$\rho_{\mathscr{D}}(x_n, \bar{x}) \leqslant \rho_{\mathscr{D}}(x_n(t), x^*(\mu_n^{-1}(t))) +$$
$$+ \rho_{\mathscr{D}}(x^*(\mu_n^{-1}(t)), \bar{x}(\varphi_n(\mu_n^{-1}(t)))) + \rho_{\mathscr{D}}(\bar{x}(t), \bar{x}(\varphi_n(\mu_n^{-1}(t)))) \leqslant$$
$$\leqslant \sup_t |x_n(\mu_n(t)) - x^*(t)| + \sup_t |x^*(t) - \bar{x}(\varphi_n(t))| +$$
$$+ \sup_t |t - \varphi_n(\mu_n^{-1}(t))| \leqslant \frac{1}{2^n} + \frac{3}{n} +$$
$$+ \sup_t |\mu_n(t) - \varphi_n(t)| \leqslant \frac{1}{2^n} + \frac{3}{n} + \frac{1}{2^n} + \varepsilon_n.$$

Par suite $\rho_{\mathscr{D}}(x_n, \bar{x}) \to 0$, c'est-à-dire la suite x_n converge vers la fonction $\bar{x}(t)$. ■

THÉORÈME 2. *Si les répartitions de $\xi_n(t)$, processus ne présentant pas de discontinuités de seconde espèce, tendent vers les répartitions de $\xi(t)$ et*

$$\lim_{c\to 0}\overline{\lim_{n\to\infty}}\,\mathsf{P}\{\Delta_c(\xi_n(t))>\varepsilon\}=0,\ \ \forall\varepsilon>0, \tag{6}$$

alors la répartition de $f(\xi_n(t))$ convergera vers celle de $f(\xi(t))$ pour toute fonctionnelle f définie sur $\mathscr{D}_{[0,1]}$ et continue dans la métrique $\rho_{\mathscr{D}}$.

Démonstration. La remarque du § 2 nous dit que la condition (6) entraîne la condition suivante:

$$\lim_{c\to 0}\sup_n \mathsf{P}\{\Delta_c(\xi_n(t))>\varepsilon\}=0. \tag{7}$$

Les remarques 1 et 4 du § 1 et le théorème 1 nous montrent que la démonstration du théorème passe par celle de

$$\lim_{L\to\infty}\sup_n \mathsf{P}\{\sup_t|\xi_n(t)|>L\}=0. \tag{8}$$

Or pour toute fonction $x(t)\in\mathscr{D}_{[0,1]}$ on a

$$\sup_{0\leqslant t\leqslant 1}|x(t)|\leqslant \sup_{0\leqslant k\leqslant m}\left|x\left(\frac{k}{m}\right)\right|+\Delta_{\frac{1}{m}}(x)$$

puisque pour $t\in\left[\frac{k}{m},\ \frac{k+1}{m}\right]$

ou bien

$$\left|x(t)-x\left(\frac{k}{m}\right)\right|<\Delta_{\frac{1}{m}}(x),$$

ou bien

$$\left|x(t)-x\left(\frac{k+1}{m}\right)\right|<\Delta_{\frac{1}{m}}(x).$$

Par suite

$$\mathsf{P}\{\sup_t|\xi_n(t)|>L\}\leqslant \mathsf{P}\left\{\sup_{k\leqslant m}\left|\xi_n\left(\frac{k}{m}\right)\right|>L-\varepsilon\right\}+ \\ +\mathsf{P}\{\Delta_{\frac{1}{m}}(\xi_n(\cdot))>\varepsilon\}.$$

La variable aléatoire $\sup_{k\leqslant m}\left|\xi_n\left(\frac{k}{m}\right)\right|$ est stochastiquement bornée uniformément en n. Ceci est une conséquence de la convergence des répartitions de $\xi_n(t)$ vers celles de $\xi(t)$. Donc les répartitions de $\sup_{k\leqslant m}\left|\xi_n\left(\frac{k}{m}\right)\right|$ convergent vers celles de $\sup_{k\leqslant m}\left|\xi\left(\frac{k}{m}\right)\right|$ et par suite

$$\overline{\lim_{L\to\infty}}\sup_n \mathsf{P}\{\sup_t|\xi_n(t)|>L\}\leqslant \sup_n \mathsf{P}\{\Delta_{\frac{1}{m}}(\xi_n(\cdot))>\varepsilon\}.$$

En passant à la limite pour $m\to\infty$, on obtient (8). ■

§ 6. Convergence de sommes de variables aléatoires indépendantes équiréparties vers un processus homogène à accroissements indépendants

Soient $\xi_{n1}, \xi_{n2}, \ldots, \xi_{nn}$ des variables aléatoires indépendantes équiréparties pour chaque n. Posons $S_{nk} = \xi_{n1} + \ldots + \xi_{nk}$. Soit le processus aléatoire $\xi_n(t)$ défini par $\xi_n(t) = S_{nk}$, $t \in \left[\frac{k-1}{n}, \frac{k}{n}\right[$ $(S_{n0} = 0)$, $\xi_n(1) = S_{nn}$. C'est un processus à accroissements indépendants présentant des discontinuités aux points k/n. Dans ce paragraphe nous allons étudier les conditions sous lesquelles les répartitions des processus $\xi_n(t)$ et les répartitions de fonctionnelles de ces processus, continues dans la métrique $\rho_{\mathscr{D}}$, convergeront vers les répartitions finidimensionnelles et les répartitions des fonctionnelles correspondantes d'un processus homogène $\xi(t)$ à accroissements indépendants. On admettra que les réalisations du processus $\xi(t)$ (qui est supposé être stochastiquement continu) appartiennent presque sûrement à $\mathscr{D}_{[0,1]}$. On sait (cf. chapitre I, § 3) que la fonction caractéristique d'un processus homogène à accroissements indépendants peut s'écrire

$$\mathsf{E}e^{i\lambda\xi(t)} = \exp\left\{t\left[i\lambda\gamma + \int\left(e^{i\lambda u} - 1 - \frac{i\lambda u}{1+u^2}\right)\frac{1+u^2}{u^2}\,dG(u)\right]\right\}, \tag{1}$$

où G est une fonction monotone bornée; à noter que la formule (1) définit entièrement les répartitions finidimensionnelles du processus $\xi(t)$. Associons à la suite des sommes S_{nk} les quantités:

$$\gamma_n = n\int_{-\infty}^{\infty}\frac{x}{1+x^2}\,dF_n(x), \quad G_n(x) = n\int_{-\infty}^{x}\frac{u^2}{1+u^2}\,dF_n(u), \tag{2}$$

où $F_n(x)$ est la fonction de répartition des ξ_{nk}.

Théorème 1. *Si existent un nombre γ et une fonction $G(x)$ non décroissante bornée tels que $\gamma_n \to \gamma$ et $G_n(x) \Rightarrow G(x)$, alors les répartitions finidimensionnelles des processus $\xi_n(t)$ convergeront vers celles d'un processus $\xi(t)$ de fonction caractéristique* (1). *De plus la répartition de $f(\xi_n(\cdot))$ convergera vers celle de $f(\xi(\cdot))$ pour toute fonctionnelle f définie sur $\mathscr{D}_{[0,1]}$ et continue dans la métrique $\rho_{\mathscr{D}}$.*

Démonstration. Le processus $\xi_n(t)$ étant à accroissements indépendants, pour démontrer la convergence des répartitions il suffit de prouver que la répartition de $\xi_n(t_2) - \xi_n(t_1)$ tend vers celle de $\xi(t_2) - \xi(t_1)$ pour $0 \leqslant t_1 < t_2 \leqslant 1$. Or ceci découle du théorème limite pour sommes de variables aléatoires indépendantes (cf. B. Gnédenko [3]). La remarque 4, § 1, et le théorème 2, § 5, nous indiquent que pour démontrer le théorème il suffit de

prouver que pour tout $\varepsilon > 0$

$$\lim_{c\to 0} \overline{\lim_{n\to\infty}} \mathsf{P}\{\Delta_c(\xi_n(\cdot)) > \varepsilon\} = 0. \tag{3}$$

Pour montrer (3) on aura besoin du lemme suivant.

LEMME 1. *Si* $\xi_1, \xi_2, \ldots, \xi_n$ *sont des variables aléatoires indépendantes équiréparties, alors*

$$\mathsf{P}\left\{\sup_{0\leqslant i<j<l\leqslant n} \min\left[\left|\sum_{k=i+1}^{j} \xi_k\right|; \left|\sum_{k=j+1}^{l} \xi_k\right|\right] > \varepsilon\right\} \leqslant$$

$$\leqslant \left(\mathsf{P}\left\{\sup_k \left|\sum_{i=1}^{k} \xi_i\right| > \frac{\varepsilon}{2}\right\}\right)^2. \tag{4}$$

Démonstration. Soit

$$i < j < l \text{ et } \left|\sum_{k=i+1}^{j} \xi_k\right| > \varepsilon, \quad \left|\sum_{k=j+1}^{l} \xi_k\right| > \varepsilon.$$

Alors, ou

$$\left|\sum_{k=1}^{i} \xi_k\right| > \frac{\varepsilon}{2},$$

ou bien

$$\left|\sum_{k=1}^{j} \xi_k\right| > \frac{\varepsilon}{2}, \quad \forall r < j,$$

ou

$$\left|\sum_{k=r+1}^{j} \xi_k\right| > \frac{\varepsilon}{2},$$

ou enfin

$$\left|\sum_{k=r+1}^{l} \xi_k\right| > \frac{\varepsilon}{2}.$$

Donc l'événement

$$\left\{\sup_{i<j<l} \min\left[\left|\sum_{k=i+1}^{j} \xi_k\right|; \left|\sum_{k=j+1}^{l} \xi_k\right|\right] > \varepsilon\right\}$$

implique l'un des événements A_r $(r = 1, \ldots, n)$:

$$A_r = \left\{|\xi_1 + \ldots + \xi_k| \leqslant \frac{\varepsilon}{2},\ k \leqslant r-1;\ |\xi_1 + \ldots + \xi_r| > \frac{\varepsilon}{2};\right.$$

$$\left.\sup_{l>r} |\xi_{r+1} + \ldots + \xi_l| > \frac{\varepsilon}{2}\right\}.$$

Comme

$$\mathsf{P}(A_r)=\mathsf{P}\left\{|\xi_1+\ldots+\xi_k|\leqslant\frac{\varepsilon}{2},\ k\leqslant r-1;\ |\xi_1+\ldots+\xi_r>\frac{\varepsilon}{2}\right\}\times$$
$$\times\mathsf{P}\left\{\sup_{l>r}|\xi_{r+1}+\ldots+\xi_l|>\frac{\varepsilon}{2}\right\},$$

l'équirépartition des ξ_i donne

$$\mathsf{P}\left\{\sup_{i<j<l}\min\left[\left|\sum_{k=i+1}^{j}\xi_k\right|;\ \left|\sum_{k=j+1}^{l}\xi_k'\right|\right]>\varepsilon\right\}\leqslant$$
$$\leqslant\sum_r\mathsf{P}\left\{|\xi_1+\ldots+\xi_k|\leqslant\frac{\varepsilon}{2},\ k\leqslant r-1;\ |\xi_1+\ldots+\xi_r|>\frac{\varepsilon}{2}\right\}\times$$
$$\times\mathsf{P}\left\{\sup_{l>r}|\xi_{r+1}+\ldots+\xi_l|>\frac{\varepsilon}{2}\right\}\leqslant$$
$$\leqslant\mathsf{P}\left\{\sup_{1\leqslant l\leqslant n}|\xi_1+\ldots+\xi_l|>\frac{\varepsilon}{2}\right\}\sum_r\mathsf{P}\left\{|\xi_1+\ldots+\xi_k|\leqslant\frac{\varepsilon}{2},\right.$$
$$\left.k\leqslant r-1;\ |\xi_1+\ldots+\xi_r|>\frac{\varepsilon}{2}\right\}=$$
$$=\left(\mathsf{P}\left\{\sup_{1\leqslant l\leqslant n}|\xi_1+\ldots+\xi_l|>\frac{\varepsilon}{2}\right\}\right)^2.\ \blacksquare$$

Reprenons la démonstration du théorème. On remarquera que

$$\Delta_c(x(\cdot))\leqslant\sup_{0\leqslant t\leqslant c}|x(t)-x(0)|+\sup_{1-c\leqslant t\leqslant 1}|x(t)-x(1)|+$$
$$+\max_{0<k<\frac{1}{c}}\ \sup_{kc\leqslant t'<t''<t'''\leqslant(k+3)c}\min\left[|x(t')-x(t'')|;\ |x(t'')-x(t''')|\right].$$

Par suite

$$\mathsf{P}\{\Delta_c(\xi_n(\cdot))>\varepsilon\}\leqslant\mathsf{P}\left\{\sup_{0\leqslant t\leqslant c}|\xi_n(t)-\xi_n(0)|>\frac{\varepsilon}{4}\right\}+$$
$$+\mathsf{P}\left\{\sup_{1-c\leqslant t\leqslant 1}|\xi_n(t)-\xi_n(1)|>\frac{\varepsilon}{4}\right\}+$$
$$+\sum_{k<\frac{1}{c}}\mathsf{P}\left\{\sup_{kc\leqslant t'<t''<t'''\leqslant(k+3)c}\min\left[|\xi_n(t'')-\xi_n(t')|;\right.\right.$$
$$\left.\left.|\xi_n(t''')-\xi_n(t'')|\right]>\frac{\varepsilon}{2}\right\}\leqslant\mathsf{P}\left\{\sup_{0\leqslant t\leqslant c}|\xi_n(t)-\xi_n(0)|>\frac{\varepsilon}{4}\right\}+$$

$$+\mathsf{P}\left\{\sup_{1-c\leqslant t\leqslant 1}|\xi_n(1)-\xi_n(t)|>\frac{\varepsilon}{4}\right\}+$$

$$+\sum_{k<\frac{1}{c}}\left(\mathsf{P}\left\{\sup_{kc\leqslant t\leqslant (k+3)c}|\xi_n(t)-\xi_n(kc)|>\frac{\varepsilon}{4}\right\}\right)^2.$$

Si $n<1/c$, il est immédiat que $\Delta_c(\xi_n(\cdot))=0$, car $\xi_n(t)$ est constant sur $\left[\frac{i}{n}, \frac{i+1}{n}\right[$.

Si $n>1/c$, même si le nombre de points de la forme i/n de l'intervalle $[kc, (k+3)c]$ varie avec k, il ne sera pas supérieur à celui des points de $[0, 4c]$ de la même forme, de sorte que toujours

$$\mathsf{P}\{\Delta_c(\xi_n(\cdot))>\varepsilon\}\leqslant 2\mathsf{P}\{\sup_{0\leqslant t\leqslant 4c}|\xi_n(t)-\xi_n(0)|>\varepsilon/4\}+$$

$$+\left(1+\frac{1}{c}\right)\left(\mathsf{P}\left\{\sup_{0\leqslant t\leqslant 4c}|\xi_n(t)-\xi_n(0)|>\frac{\varepsilon}{4}\right\}\right)^2.$$

Pour évaluer la probabilité

$$\mathsf{P}\left\{\sup_{0\leqslant t\leqslant 4c}|\xi_n(t)-\xi_n(0)|>\frac{\varepsilon}{4}\right\}=\mathsf{P}\left\{\sup_{k\leqslant 4nc}\left|\sum_{i=1}^{k}\xi_{ni}\right|>\frac{\varepsilon}{4}\right\}$$

on se servira des quantités: $\xi'_{ni}=\xi_{ni}$ si $|\xi_{ni}|\leqslant L$, $\xi'_{ni}=0$ si $|\xi_{ni}|>L$, $\xi''_{ni}=\xi_{ni}-\xi'_{ni}$. Il vient

$$\mathsf{P}\left\{\sup_{1\leqslant k\leqslant N}\left|\sum_{i=1}^{k}\xi_{ni}\right|>\frac{\varepsilon}{4}\right\}\leqslant\mathsf{P}\{\sup_{1\leqslant k\leqslant N}|\xi''_{nk}|>0\}+$$

$$+\mathsf{P}\left\{\sup_{1\leqslant k\leqslant N}\left|\sum_{i=1}^{k}\xi'_{ni}\right|>\frac{\varepsilon}{4}\right\}=\sum_{i=1}^{N}\mathsf{P}\{|\xi_{ni}|>L\}+$$

$$+\mathsf{P}\left\{\sup_{1\leqslant k\leqslant N}\left|\sum_{i=1}^{k}(\xi'_{ni}-\mathsf{E}\xi'_{ni})\right|>\frac{\varepsilon}{4}-\left|\sum_{i=1}^{N}\mathsf{E}\xi'_{ni}\right|\right\}\leqslant$$

$$\leqslant N\mathsf{P}\{|\xi_{n1}|>L\}+\frac{N\operatorname{Var}\xi'_{n1}}{\left(\frac{\varepsilon}{4}-N|\mathsf{E}\xi'_{n1}|\right)^2}$$

(nous avons utilisé l'inégalité de Kolmogorov qui constitue la remarque du théorème 5, § 1, chapitre III), si seulement $N|\mathsf{E}\xi'_{n1}|<\varepsilon/4$.

Si L et $-L$ sont des points de continuité de la fonction G, alors

$$\lim_{n\to\infty} n\mathsf{P}\{|\xi_{n1}|>L\}=\lim_{n\to\infty}\int_{|x|>L}\frac{1+x^2}{x^2}\,dG_n(x)=\int_{|x|>L}\frac{1+x^2}{x^2}\,dG(x),$$

$$\lim_{n\to\infty} n\mathsf{E}\xi'_{n1}=\lim_{n\to\infty} n\int_{|x|<L} x\,dF_n(x)=$$

$$=\gamma-\lim_{n\to\infty} n\int_{|x|>L}\frac{x}{1+x^2}\,dF_n(x)+\lim_{n\to\infty} n\int_{|x|\leqslant L}\left(x-\frac{x}{1+x^2}\right)dF_n(x)=$$

$$=\gamma-\int_{|x|>L}\frac{1}{x}\,dG(x)+\int_{|x|\leqslant L} x\,dG(x)=\gamma_L,$$

$$\overline{\lim_{n\to\infty}}\, n\operatorname{Var}\xi'_{n1}\leqslant\lim_{n\to\infty} n\mathsf{E}(\xi'_{n1})^2\leqslant\lim_{n\to\infty} n\mathsf{E}\,\frac{1+L^2}{1+\xi_{n1}^2}\,\xi_{n1}^2=$$

$$=(1+L^2)\int dG(x).$$

Donc si $4c|\gamma_L|<\varepsilon/4$, alors

$$\overline{\lim_{n\to\infty}}\,\mathsf{P}\left\{\sup_{1\leqslant k\leqslant 4nc}\left|\sum_{i=1}^{k}\xi_{ni}\right|>\frac{\varepsilon}{4}\right\}\leqslant 4c\int_{|x|>L}\frac{1+x^2}{x^2}\,dG(x)+$$

$$+\frac{4c(1+L^2)\int dG(x)}{\left(\frac{\varepsilon}{4}-4c|\gamma_L|\right)^2}.$$

Par suite, pour les c assez petits

$$\overline{\lim_{n\to\infty}}\,\mathsf{P}\left\{\sup_{0\leqslant t\leqslant 4c}|\xi_n(t)-\xi_n(0)|>\frac{\varepsilon}{4}\right\}\leqslant Kc,$$

où K est une constante telle que

$$\overline{\lim_{n\to\infty}}\,\mathsf{P}\left\{\sup_{0\leqslant t\leqslant 4c}|\xi_n(t)-\xi_n(0)|>\frac{\varepsilon}{4}\right\}\leqslant Kc.$$

D'où (3). ■

Les corollaires suivants sont des théorèmes limites.

COROLLAIRE 1. *Soient $a(t)$ et $b(t)$ des fonctions continues positives, α un nombre réel. Si les conditions du théorème 1 sont réunies, alors*

$$\lim_{n\to\infty}\mathsf{P}\left\{a\left(\frac{k}{n}\right)-\alpha b\left(\frac{k}{n}\right)<S_{nk}<a\left(\frac{k}{n}\right)+\alpha b\left(\frac{k}{n}\right),\ k=\overline{1,n}\right\}=$$

$$=\mathsf{P}\{a(t)-\alpha b(t)<\xi(t)<a(t)+\alpha b(t),\quad 0\leqslant t\leqslant 1\}$$

pour tous les $\alpha>0$ tels que le second membre soit fonction continue de α.

COROLLAIRE 2. *Soit* $g(x)$ *une fonction continue définie sur* $]-\infty, \infty[$; *si sont réalisées les conditions du théorème* 1, *alors*

$$\lim_{n\to\infty} \mathsf{P}\left\{\frac{1}{n}\sum_{k=1}^{n} g(S_{nk}) < \alpha\right\} = \mathsf{P}\left\{\int_0^1 g(\xi(t))\,dt < \alpha\right\}$$

pour tous les α *tels que*

$$\mathsf{P}\left\{\int_0^1 g(\xi(t))\,dt = \alpha\right\} = 0.$$

On démontre ces propositions en considérant les fonctionnelles

$$f_1(x(\cdot)) = \sup_t \frac{1}{b(t)}\,|x(t) - a(t)|,\quad f_2(x(\cdot)) = \int_0^1 g(x(t))\,dt.$$

Soit $\gamma_a(x(\cdot))$ une fonctionnelle telle que

$$\gamma_a(x(\cdot)) = \begin{cases} 0 & \text{pour } \sup\limits_{0\leqslant t\leqslant 1} x(t) < a, \\ x(\tau) - a & \text{pour } \sup\limits_{0\leqslant t\leqslant 1} x(t) \geqslant a, \end{cases}$$

où $\tau \in [0, 1]$ tel que $x(t) < a$ pour $t < \tau$ et $x(\tau) \geqslant a$. La quantité $\gamma_a(x(\cdot))$ est la valeur du premier saut de la fonction $x(\cdot)$ par dessus le niveau a.

COROLLAIRE 3. *Si* a *est tel que*

$$\mathsf{P}\{\sup_{t_1\leqslant s\leqslant t_2} \xi(s) = a\} = 0,\quad \forall t_1 < t_2\ (t_1, t_2 \in [0, 1]),$$

alors la répartition de $\gamma_a(\xi_n(\cdot))$ *converge vers celle de* $\gamma_a(\xi(\cdot))$.

Cette assertion découle de la remarque 3, § 1, et de la continuité de $\gamma_a(x(\cdot))$ sur tous les $x(\cdot)$ pour lesquels a n'est pas maximum local.

NOTICE BIBLIOGRAPHIQUE

La notice donne quelques indications bibliographiques sur les sujets traités et n'a pas l'ambition de passer en revue les principales notions de la théorie des processus aléatoires. Nous avons évité souvent les références aux œuvres originales d'accès difficile et avons porté notre choix sur des monographies plus récentes renfermant une bibliographie sur les sujets abordés.

Chapitre I

§ 1. L'étude systématique des problèmes de la théorie des processus aléatoires a été entamée dans les travaux de E. Sloutsky [1] et de A. Kolmogorov [7], [8]. Les travaux de A. Kolmogorov [7], [8] ont été déterminants dans la création de la théorie des processus aléatoires.

§ 2. Les processus gaussiens sont largement utilisés dans de nombreux problèmes de probabilité (cf. par exemple H. Cramer, M. Leadbetter [1]), dans les problèmes de statistique, de prédiction et de filtrage des processus aléatoires, dans la théorie de commande optimale des solutions d'équations différentielles perturbées par un processus aléatoire (R. Liptser, A. Chiriaev [1], K. Åstrem [1]). La généralisation du théorème limite central est due à S. Bernstein [3].

§ 3. Le processus du mouvement brownien a été étudié pour la première fois par L. Bachelier [1]. Le travail de B. Finetti [1] a servi d impulsion à l'étude des processus arbitraires à accroissements indépendants. A. Kolmogorov [4] a trouvé la fonction caractéristique d'un processus arbitraire à accroissements indépendants à moment d'ordre deux fini; la formule générale est due à P. Lévy [1].

§ 4. Les processus markoviens à temps continu (au sens large) et leurs principaux types ont été introduits dans le travail de A. Kolmogorov [8]. Les théorèmes d'existence des solutions des équations de Kolmogorov ont été étudiés pour la première fois par W. Feller [1], [2].

§ 5. Les processus stationnaires au sens large ont été introduits par A. Khintchine [3]. De nombreux travaux ont été consacrés aux oscillations à amplitudes et phases aléatoires (cf. par exemple N. Bogolioubov [1]). A. Khintchine [3] a déduit la représentation spectrale de la fonction de corrélation d'un processus stationnaire au sens large, H. Cramer [1] a généralisé cette représentation au cas multidimensionnel. La formule de la fonction de corrélation d'un champ aléatoire homogène et isotrope figure dans le travail de J. L. Schönberg [1].

Chapitre II

§ 1. L'axiomatique ensembliste de la théorie des probabilités actuellement en usage a été proposée par A. Kolmogorov en 1929. Elle est développée dans sa monographie [7]. La théorie de la mesure et de l'intégrale est exposée

dans les ouvrages de P. Halmos [1], A. Kolmogorov et S. Fomine [1] ainsi que dans les manuels de théorie des probabilités de J. Neveu [1], M. Loeve [1], P. L. Hennequin et A. Tortrat [1].

§ 2. Le principal théorème de ce paragraphe (le théorème 5) est prouvé dans le livre de A. Kolmogorov [7] (pour une famille de variables aléatoires). La démonstration du théorème 6 figure par exemple dans P. Halmos [1] et J. Neveu [1].

§ 3. La théorie des probabilités et des espérances mathématiques conditionnelles est l'œuvre de A. Kolmogorov [7]. Elle a été affinée par J. L. Doob [1].

§ 4. La loi générale de tout ou rien a été établie par A. Kolmogorov [7].

Chapitre III

§ 1. Les martingales ont été abordées par divers auteurs, mais leur théorie systématique est redevable à J. L. Doob [1]. On lui doit les inégalités fondamentales pour les martingales, le théorème d'existence de la limite, la notion de semi-martingale et d'autres résultats. Les martingales sont amplement traitées aussi par P. Meyer [1], J. Neveu [2].

§ 2. Les idées et résultats essentiels de ce paragraphe appartiennent à A. Kolmogorov et A. Khintchine [1] et à A. Kolmogorov [1]. Les séries de variables aléatoires indépendantes sont étudiées en détail dans les ouvrages de J. L. Doob [1], M. Loeve [1], A. Skorokhod [6].

§ 3. Le théorème ergodique doit son existence aux problèmes de mécanique statistique. Voir à ce sujet le livre de A. Khintchine [5]. Les premiers théorèmes ergodiques de J. Neumann et G. Birkhoff ont imprimé une forte impulsion à la théorie. L'ouvrage de E. Hopf [1] passe en revue les premiers résultats de la théorie ergodique. A. Kolmogorov [9] propose une démonstration simple du théorème de Birkhoff-Khintchine. La théorie ergodique est développée dans les ouvrages de P. Halmos [2], P. Billingsley [1].

§ 4. Les problèmes de la théorie de renouvellement ont à maintes reprises fait l'objet de travaux probabilistiques appliqués. Voir W. Feller [6].

§ 5. Les chaînes de Markov à nombre fini d'états ont été conçues (en 1906) et étudiées par A. Markov [1]. La définition générale d'une chaîne et d un processus de Markov revient à A. Kolmogorov [8].

§ 6. Les chaînes de Markov à nombre dénombrable d'états apparaissent pour la première fois dans les travaux de A. Kolmogorov [6]. On les retrouve chez d'autres auteurs. Voir W. Feller [3], Chung K. L. [1], J. G. Kemeny, J. L. Snell, A. W. Knapp [1].

Chapitre IV

§§ 1 à 3. E. Sloutsky et A. Kolmogorov se sont penchés les premiers sur la possibilité de construire un processus aléatoire stochastiquement équivalent à un processus donné, dont les réalisations vérifient certaines conditions de régularité (voir E. Sloutsky [2]). J. L. Doob a fortement contribué à l'élaboration des problèmes et des diverses variantes de la définition axiomatique de la fonction aléatoire. On lui doit les idées et les théorèmes principaux des §§ 2, 3.

§ 4. Une formulation moins forte du théorème 1 a été prouvée par N. Tchentsov [1], le théorème 2, par J. H. Kinney [1] (pour les processus markoviens). L'absence de discontinuités de seconde espèce pour les processus stochastiquement continus à accroissements indépendants a été établie par P. Levy [1].

§ 5. Le théorème 1 a été démontré par E. Dynkine [1] et J. H. Kinney [1] (pour les processus markoviens) indépendamment l'un de l'autre. Une variante plus faible du théorème 6 a été proposée par A. Kolmogorov et publiée pour la première fois dans E. Sloutsky [2]; pour les propriétés locales des processus gaussiens on pourra consulter H. Cramer et M. Leadbetter [1].

§ 6. Les propriétés des réalisations des semi-martingales ont été étudiées par J. L. Doob [1].

Chapitre V

§ 1. Une introduction à la théorie de l'espace hilbertien figure dans A. Kolmogorov et S. Fomine [1]. On trouvera un exposé complet dans N. Ahiezer et I. Glazman [1].

§ 2. Voir E. Sloutsky [1], M. Loeve [1], [2].

§ 3. La théorie des intégrales stochastiques est l'œuvre de H. Cramer [1]; A. Kolmogorov a le premier fait la liaison entre les intégrales stochastiques et la théorie des représentations spectrales des fonctions aléatoires d'un côté et la théorie de l'espace hilbertien de l'autre [10], [11], [12].

§ 4. Le théorème 1 a été démontré par K. Karhunen [1], le théorème 2 par H. Cramer [1]. Pour une plus ample connaissance de la théorie des processus stationnaires il est recommandé de voir E. J. Hannan [1], G. M. Jenkins, D. G. Watts [1].

§ 5. Une théorie plus générale des transformations linéaires des processus aléatoires peut être bâtie à l'aide de la théorie des processus aléatoires généraux, proposée par I. Guelfand et K. Ito (voir I. Guelfand et N. Vilenkine [1], K. Ito [5]).

§ 6. Des résultats fondamentaux sur les suites stationnaires ont été obtenus par A. Kolmogorov [12], sur les processus stationnaires par K. Karhunen [2].

Chapitre VI

§ 1. Dans le livre de F. Spitzer [2] on trouvera un exposé de la théorie générale des errements, ainsi que des références aux sources. F. Spitzer [1] a trouvé les conditions de majoration d'un errement et les formules de répartition du maximum. B. Rogozine [1] s'est penché sur la répartition de la valeur et de l'instant du saut.

§ 2. Le processus général de Poisson a été introduit par A. Khintchine [1]. La répartition de la valeur et de l'instant de saut par dessus un niveau fait l'objet des travaux de B. Rogozine [2] et D. Goussak [1].

§ 3. N. Wiener [1] a construit de façon rigoureuse le processus wienérien et a étudié les propriétés de ses réalisations. La condition de continuité d'un processus à accroissements indépendants a été établie par A. Khintchine [1]. Les résultats des théorèmes 2 et 3 découlent des résultats généraux de I. Pétrovski [1].

§ 4. P. Levy [1] propose une décomposition d'un processus en une composante continue et une composante discontinue. On trouve encore dans ce paragraphe la formule générale de la fonction caractéristique d'un processus à accroissements indépendants. Dans certains cas particuliers cette formule a été déduite par B. Finetti [1] et A. Kolmogorov [4]. La méthode d'étude d'un processus à l'aide d'une mesure construite par sauts a été proposée par K. Ito [1], [7].

§ 5. La croissance des processus homogènes à accroissements indépendants a été abordée par A. Khintchine [4], B. Gnédenko [1], [2]. La loi du logarithme itéré pour un processus wienérien a été prouvée par A. Khintchine [1].

Chapitre VII

§ 1. Le travail de A. Kolmogorov [8] constitue l'ossature de la théorie générale des processus markoviens. J. L. Doob [1] a procédé par la suite à une analyse de la définition du processus markovien. La définition la plus générale de cette notion figure dans E. Dynkine [4]. La propriété d'être markovien fort a été étudiée par J. L. Doob [1], E. Dynkine [5], E. Dynkine et A. Youchkévitch [1]. Une condition suffisante d'être markovien fort a été établie par E. Dynkine et A. Youchkévitch [1].

§ 2. J. L. Doob s'est penché sur les processus discontinus à espace des phases arbitraire et a exposé ses résultats dans son livre [1].

§ 3. Les processus à nombre dénombrable d'états ainsi que la déduction des équations de Kolmogorov sont traités dans A. Kolmogorov [8]. La dérivabilité de la probabilité de passage a été établie par A. Kolmogorov [13]. Les théorèmes d'existence des solutions des équations de Kolmogorov ont été étudiés par W. Feller [1], [2]. Le livre de Chung K. L. [1] est consacré à la théorie générale des processus homogènes à nombre dénombrable d'états.

§ 4. L'ouvrage de W. Feller [3] renferme d'innombrables exemples de processus markoviens, y compris les processus de naissance et de mort.

§ 5. H. M. Watson et W. Galton [1] ont les premiers considéré les processus branchus. La définition générale du processus branchu fait l'objet d'un article de A. Kolmogorov et N. Dmitriev [1]. Ce paragraphe est basé sur un aperçu de B. Sébastianov [1].

Chapitre VIII

L'interprétation probabiliste du phénomène de diffusion a été envisagée par A. Khintchine [1]. Les équations différentielles stochastiques pour processus aléatoires ont été étudiées par S. Bernstein [1], I. Guikhman [1], [2], K. Ito [3], [4]. On se sert essentiellement de la terminologie et des notations de Ito. Pour un exposé plus général de la théorie des équations différentielles stochastiques le lecteur est renvoyé à I. Guikhman et A. Skorokhod [1].

§ 1. Les principaux résultats de ce paragraphe sont dus à K. Ito [2], [6].

§ 2. Dans cette forme les équations ont été étudiées par K. Ito [3], [4]; ce dernier a prouvé le théorème d'existence et d'unicité et aussi que la solution sera processus markovien.

§ 3. La dérivabilité des solutions des équations stochastiques par rapport aux conditions initiales a été établie par I. Guikhman [2].

§ 4. L idée de déduire les équations de Kolmogorov en se servant de la dérivabilité de la solution d'une équation stochastique par rapport aux conditions initiales a été avancée par I. Guikhman [2]. La déduction des équations pour la répartition d'une fonctionnelle additive d'un processus du mouvement brownien revient à M. Kac [1], [2] et dans le cas général à E. Dynkine [3].

§ 5. I. Pétrovski [1] a proposé d'appliquer les équations différentielles aux errements dans un domaine borné. Les processus de diffusion ont été examinés dans des domaines bornés par A. Khintchine [1]. R. Kasminski [1] a envisagé les répartitions de fonctionnelles liées avec le temps d'atteinte des frontières par un processus de diffusion à une dimension. Les processus de diffusion à une dimension sont étudiés dans l'ouvrage de I. Guikhman et A. Skorokhod [1].

§ 6. I. Guirsanov [1] et A. Skorokhod [4] ont étudié les conditions de continuité absolue des mesures et la forme de la densité pour les processus de diffusion.

Chapitre IX

A. Kolmogorov [3], [5], I. Pétrovski [1] et A. Khintchine [1] ont les premiers considéré les théorèmes limites pour les probabilités d'événements dépendant de la trajectoire tout entière du processus (la probabilité que le processus reste dans une bande curviligne). Le premier théorème limite général pour des fonctionnelles quelconques continues dans la métrique de $\mathscr{C}$ a été obtenu par M. Donsker [1] (dans le cas où les sommes des variables aléatoires indépendantes équiréparties convergent vers un processus du mouvement brownien).

§ 1. La faible convergence des mesures dans des espaces métriques a été traitée par You. Prokhorov [1], [3].

§ 2. You. Prokhorov [1], [3] a prouvé un théorème limite général pour processus continus.

§ 3. A. Kolmogorov [3], [5], M. Kac et P. Erdös [1], [2] ont envisagé les théorèmes limites pour divers cas particuliers. Le théorème général est dû à You. Prokhorov [1], [3], le cas particulier (corollaire) a été établi antérieurement par M. Donsker [1].

§ 4. S. Bernstein [1], [3], A. Khintchine [1], I. Guikhman [4], [5] ont étudié des cas concrets de convergence vers des processus de diffusion, H. Maruyama [1], You. Prokhorov [2], A. Skorokhod [4], [5] des théorèmes pour des fonctionnelles $\mathscr{C}$-continues.

§ 5. La convergence traitée dans ce paragraphe a été introduite par A. Skorokhod. N. Tchentsov a établi un théorème limite intéressant pour processus sans discontinuités de seconde espèce.

§ 6. I. Guikhman [3] a établi un théorème limite pour la probabilité qu'une suite de sommes reste dans une bande curviligne. A. Skorokhod [2], You. Prokhorov [3] ont considéré des théorèmes généraux pour fonctionnelles.

BIBLIOGRAPHIE

AHIEZER N., GLAZMAN I.
[1] Théorie des opérateurs linéaires dans un espace hilbertien (en russe), Moscou, 1966.

ÅSTREM K.
[1] Introduction to stochastic control theory, N.Y., 1970.

BACHELIER L.
[1] Théorie de la spéculation, Ann. Sci. Ecol. Norm. Sup. 3 (1900), 21-86.

BERNSTEIN S.
[1] Principes de la théorie des équations différentielles stochastiques. Troudy Fiz.-matem. in-ta im. Stéklova 5 (1934), 95-124.
[2] Extension d'un théorème limite de la théorie des probabilités à une somme de variables liées, Ouspekhi matem. naouk 10 (1944), 65-114.
[3] Théorie des probabilités (en russe), Gostekhizdat, 1946.

BILLINGSLEY P.
[1] Ergodic theory and information, N.Y., 1965.

BOCHNER S.
[1] Harmonic analysis and the theorie of probability, Berkeley and Los Angeles, 1955.

BOGOLIOUBOV N.
[1] Sur certaines méthodes de statistique en physique mathématique (en russe), Izd-vo AN USSR, 1945.

CHUNG K. L.
[1] Markov chains with stationary transition probabilities, W. Berlin, 1960.
[1] On the theorie of random processes, Ann. Math. 41 (1940), 215-230.

CRAMER H.
[1] On the theory of random processes, Ann. math. 41 (1940), 215-230.

CRAMER H., LEADBETTER M.
[1] Stationary and related stochastic processes, N.Y., 1967.

DONSKER M.
[1] An invariance principle for certain probability limit theorems, Mem. Amer. Math. Soc. 6 (1951), 1-12.
[2] Justification and extension of Doob's heuristic approach to the Kolmogoroff-Smirnov theorems, Ann. Math. Stat. 23 (1952), 277-281.

DOOB J. L.
[1] Stochastic processes. N.Y., 1953.

DYNKINE E.
[1] Critère de continuité et d'absence de discontinuités de seconde espèce pour les trajectoires d'un processus aléatoire markovien (en russe), Izv. AN SSSR, ser. matem., 16 (1952), 563-572.
[2] Quelques théorèmes limites pour sommes de variables aléatoires indépendantes d'espérances mathématiques infinies (en russe), Izv. AN SSSR, ser. matem., 19 (1955), 247-266.

[3] Fonctionnelles de trajectoires de processus aléatoires markoviens (en russe), Doklady AN SSSR **104** (1955), 691-694.
[4] Eléments de la théorie des processus markoviens (en russe), Fizmatguiz, 1959.
[5] Processus markoviens (en russe), Fizmatguiz, 1963.

DYNKINE E., YOUCHKÉVITCH A.
[1] Processus markoviens forts (en russe), Theoria veroïatnosteï i eïo primenenia **1** (1956), 149-155.

EDWARDS R.
[1] Functional analysis. Theory and applications, N.Y., 1965.

EINSTEIN A., SMOLUCHOWSKI M.
[1] Brownian motion, ONTI, 1936.

FELLER W.
[1] Zur Theorie der stochastischen Prozesse, Math. Ann. **113** (1936), 113-160.
[2] On integro-differential equations of purely discontinuous Markov processes, Trans. Amer. Math. Soc. **46** (1940), 488-515.
[3] An introduction to probability theory and its applications, Vol. 1, N.Y., 1966.
[4] Diffusion processes in one dimension, Trans. Amer. Math. Soc. **77** (1954), 1-31.
[5] The general diffusion operator and positivity preserving semi-groups in one dimension, Ann. Match. **60** (1954), 417-435.
[6] An introduction to probability theory and its applications, Vol. 2, N.Y., 1966.

FINETTI B.
[1] Sulle funzioni a incremento aleatorio, Rend. Accad. Naz. Lincei, Cl. Sci. Fis.-Mat. Nat. (6), **10** (1929), 163-168.

GNÉDENKO B.
[1] Sur la croissance des processus aléatoires homogènes à accroissements indépendants (en russe), Izv. AN SSSR, ser. matem., **7** (1943), 89-110.
[2] Contribution à la théorie de la croissance des processus aléatoires homogènes à accroissements indépendants (en russe), Sb. troudov In-ta matem. AN SSSR **10** (1948), 60-82.
[3] Théorie des probabilités, Dunod, Paris, 1962.

GNÉDENKO B., KOLMOGOROV A.
[1] Limit distributions for series of independant random variables, Addison-Wesley Mathematics series, 1954.

GOUSSAK D.
[1] Sur la répartition conjointe de l'instant et de la valeur du premier saut de processus homogènes à accroissements indépendants (en russe), Theoria veroïatnosteï i eïo primenenia **14** (1969), 15-23.

GRENANDER W., SZEGÖ G.
[1] Toeplitz forms and their applications, Los Angeles, 1958.

GUELFAND I.
[1] Processus aléatoires généraux (en russe), Dokl. AN SSSR **100** (1955), 853-856.

GUELFAND I., VILENKINE N.
[1] Quelques applications de l'analyse harmonique. Espaces hilbertiens munis (en russe), Fizmatguiz, 1961.

GUIKHMAN I.
[1] Sur certaines équations différentielles à fonctions aléatoires (en russe), Oukr. matem. journ. **2**, n° 3 (1950), 45-69.
[2] Contribution à la théorie des équations différentielles des processus aléatoires (en russe), Oukr. matem. journ. **2**, n° 4 (1950), 37-63, **3** (1951), 317-339.
[3] Sur un théorème de A. Kolmogorov (en russe), Naoutch. zap. Kievsk. univ.. Matem. sb. **7** (1958), 76-94.

[4] Sur quelques théorèmes limites pour répartitions conditionnelles et les problèmes de statistique qui y sont rattachés (en russe), Oukr. matem. journ. 5 (1953), 413-433.
[5] Processus markoviens dans les problèmes de statistique mathématique (en russe), Oukr. matem. journ. 6 (1954), 28-36.

GUIKHMAN I., SKOROKHOD A.
[1] Stochastic differential equations, Springer-Verlag, 1972.

GUIRSANOV I.
[1] Sur les transformations d'une classe de processus aléatoires par un changement de mesure absolument continu (en russe), Theoria veroïatnosteï i eïo primenenia 5 (1960), 314-330.

HALMOS P. R.
[1] Measure theory, D. Van Nostrand, N.Y., 1950.
[2] A Hilbert space problem book, London, 1967.

HANNAN E. J.
[1] Group representations and applied probability, London, 1965.

HARRIS E. T.
[1] Some mathematical models for branching processes, Proc. II Berkeley Symp. on Math. Stat. and Probab., 1951, 305-328.
[2] The theory of branching processes, Springer Verlag., 1963.

HENNEQUIN P., TORTRAT A.
[1] Théorie des probabilités et quelques applications, Paris, Masson et Cie, 1965.

HOPF E.
1] Ergodentheorie, Ergebnisse der mathematik und ihrer Genzgebiete, New York, Chelsea Publishing Co., 1948.

IBRAHIMOV I., ROSANOV You.
[1] Processus aléatoires gaussiens, « Mir », 1974.

IRJINA M.
[1] Comportement asymptotique des processus aléatoires branchus, Journal mathématique de Tchécoslovaquie 7 (1957), 130-153.

ITO K.
[1] On stochastic processes, Jap. J. Math. 18 (1942), 261-301.
[2] Stochastic integral, Proc. Imp. Acad. Tokyo 20 (1944), 519-524.
[3] On stochastic integral equation, Proc. Jap. Acad. 1-4 (1946), 32-35.
[4] On stochastic differential equation in a differentiable manifold, Nagoya Math. J., 1 (1950), 35-47.
[5] Stationary random distribution, Mem. Coll. Sci. Univ. Kyoto 28 (1954), 209-223.
[6] On a formula concerning stochastic differentials, Nagoya Math. J., 3 (1951), 55-65.
[7] Kakuritsu Katei [Probability processes], Tokyo, Iwanami Shoten Publishers. Vol. I, 1960, Vol. II, 1963.

JENKINS G. M., WATTS D. G.
[1] Spectral analysis and its applications, San Francisco, 1966.

KAC M.
[1] On distributions of certain functionals, Trans. Amer. Math. Soc. 65 (1949), 1-13.
[2] One some connections between probability theory and differential and integral equations, Proc. Sec. Berkeley Symp. Math. Stat. and Prob., Berkeley, 1951, pp. 189-213.

KAC M., ERDÖS P.
[1] On certain limit theorems of theory of probability, Bull. Amer. Math. Soc. 53 (1946), 292-302.
[2] On the number of positive sums of independent random variables, Bull. Amer. Math. Soc. 53 (1946), 1011-1020.

KARHUNEN K.
[1] Ueber lineare Methoden in der Wahrscheinlichkeitsrechnung, Ann. Acad. Sci. Fennicae, Ser. A, Math., Phys., **37** (1947), 3-79.
[2] Ueber die Structur stationären zufälliger Functionen, Ark. Math. **1** (1950), 141-160.
KEMENY J. G., SNELL J. L., KNAPP A. W.
[1] Denumerable Markov chains, D. van Nostrand Co., 1966.
KHASMINSKI R.
[1] Répartition des probabilités pour une fonctionnelle de trajectoire d'un processus aléatoire du type de diffusion (en russe), Dokl. AN SSSR **104** (1955), 22-25.
KHINTCHINE A.
[1] Lois asymptotiques de la théorie des probabilités (en russe), ONTI, 1936.
[2] Zur Theorie der unbeschränkt teilbaren Verteilungsetze, Matem. zb. **2** (1937), 79-120.
[3] Théorie de la corrélation des processus aléatoires stationnaires (en russe), Ouspekhi matem. naouk **5** (1938), 42-51.
[4] Sur la croissance locale de processus stochastiques sans postaction (en russe), Izv. AN SSSR, ser. matem., (1939), 487-508.
[5] Fondements mathématiques de la mécanique statistique (en russe), Gostekhizdat, 1943.
KINNEY J. H.
[1] Continuity properties of simple functions of Markov processes, Trans. Amer. Math. Soc. **74** (1953), 280-302.
KOLMOGOROV A.
[1] Ueber die Summen durch den Zufall bestimmer unabhängiger Grössen, Math. Ann. **99** (1928), 309-319; **100** (1929), 484-488.
[2] Théorie générale de la mesure et calcul des probabilités (en russe), Troudy Komm. acad., sect. matem. **1** (1929), 8-21.
[3] Eine Verallgemeinerung des Laplace-Liapounoffschen Satzes, Izvestia AN SSSR, Otdelenie matem. i estestv. naouk (1931), 959-962.
[4] Sulla forma generale di un processo stocastico omogeneo, Atti Acad. Lincei **15** (1932), 805-808, 866-869.
[5] Ueber die Grenzwertsätze der Wahrscheinlichkeitsrechnung, Izvestia AN SSSR, Otdelenie matem. i estesv. naouk (1933), 363-372.
[6] Chaînes de Markov à nombre dénombrables d'états (en russe), Bull. MGU **1**, n° 3 (1937), 1-16.
[7] Fundations of the theory of Probability. Chelsea Publishing Company, N.Y., 1956.
[8] Sur les méthodes analytiques en théorie des probabilités (en russe), Ouspekhi matem. naouk **5** (1938), 5-41.
[9] Démonstration simplifiée du théorème ergodique de Birkhoff-Khintchine (en russe), Ouspekhi matem. naouk **5** (1938), 52-56.
[10] Courbes d'un espace hilbertien invariantes par rapport à un groupe de mouvements à un paramètre (en russe), Dokl. AN SSSR **26** (1940), 6-9.
[11] Spirale de Wiener et autres courbes intéressantes dans un espace hilbertien (en russe), Dokl. AN SSSR **26** (1940), 115-118.
[12] Suites stationnaires dans un espace hilbertien, Bull. MGU **2**, n° 6 (1941), 1-40.
[13] Sur la dérivabilité des probabilités de passage de processus markoviens homogènes dans le temps à nombre dénombrable d'états (en russe), Outchenye zap. MGU, ser. matem., **4**, vyp. 148 (1951), 53-59.
KOLMOGOROV A., DMITRIEV N.
[1] Processus aléatoires branchus (en russe), Dokl. AN SSSR **56** (1947), 7-10.
KOLMOGOROV A., FOMINE S.
[1] Eléments de la théorie des fonctions et de l'analyse fonctionnelle, « Mir », 1977.

KOLMOGOROV A., KHINTCHINE A.
[1] Über Konvergenz von Rhein deren Glieder durch den Zufalle bestimmet werden. Matem. zb. 32 (1952), 668-677.
KREÏN M.
[1] Sur un problème d'interpolation de A. Kolmogorov (en russe), Dokl. AN SSSR 46 (1944), 306-309.
[2] Sur le problème fondamental d'approximation de la théorie d'extrapolation et de filtrage des processus aléatoires stationnaires, Dokl. AN SSSR 94 (1954), 13-16.
LEVY P.
[1] Sur les intégrales dont les éléments sont des variables aléatoires indépendantes, Ann. Scuola Norm., Pisa 2, nº 3 (1934), 337-366.
[2] Sur certains processus stochastiques homogènes, Comp. Math. 7 (1939), 283-339.
[3] Processus stochastiques et mouvement brownien, Paris, 1965.
LIPTSER R., CHIRIAEV A.
[1] Statistique des processus aléatoires (en russe), « Naouka », 1974.
LOEVE M.
[1] Fonctions aléatoires du second ordre, note dans P. Levy [3].
[2] Probability theory, Princeton, 1960.
MARKOV A.
[1] Extension de la loi des grands nombres à des quantités dépendant l'une de l'autre (en russe), Izv. Fiz-matem. o-va pri Kazanskom un-te (2) 15 (1906), 135-156.
MARUYAMA G.
[1] Continious Markov processes and stochastic equations. Rend. Circolo Math. Palermo, 4 (1955), 1-43.
MEYER Paul A.
[1] Probabilités et potentiels. Paris, Hermann, 1966, Publications de l'Institut de Mathématique de l'Université de Strasbourg, nº XIV.
NEVEU J.
[1] Bases mathématiques du calcul des probabilités. Paris, Masson et Cie, 1964.
[2] Martingales à temps discret. Masson et Cie, Paris, 1972.
PÉTROV V.
[1] Sommes de variables aléatoires indépendantes (en russe), « Naouka », 1972.
PÉTROVSKY I.
[1] Ueber das Irrfahrproblem, Math. Ann. 109 (1934), 425-444.
POUGATCHEV V.
[1] Théorie des fonctions aléatoires et applications aux problèmes de commande automatique (en russe), Gostekhizdat, 1957.
PRIVALOV I.
[1] Propriétés frontières des fonctions analytiques (en russe), Gostekhizdat, 1950.
PROKHOROV You.
[1] Répartition des probabilités dans des espaces fonctionnels (en russe), Ouspekhi matem. naouk 8, nº 3 (1953), 165-167.
[2] Méthodes d'analyse fonctionnelle dans les théorèmes limites de la théorie des probabilités (en russe), Vestn. Leningrad. un-ta 11 (1954), 44.
[3] Convergence des processus aléatoires et des théorèmes limites de la théorie des probabilités (en russe), Theoria veroïatnosteï i eïo primenenia 1 (1956), 177-238.
PROKHOROV You, ROZANOV You.
[1] Théorie des probabilités (en russe), « Naouka », 1973.
REUTER G. E. H.
[1] Denumerable Markov processes and the associated contraction semigroups on L, Acta Math. 97 (1957), 1-46.

ROGOZINE B.
[1] Répartition de la valeur du premier saut (en russe), Theoria veroïatnosteï i eïo primenenia 9 (1964), 498-515.
[2] Sur la répartition de certaines fonctionnelles rattachées aux problèmes aux limites pour processus à accroissements indépendants (en russe), Theoria veroïatnosteï i eïo primenenia 11 (1966), 656-670.
ROZANOV You.
[1] Processus aléatoires, Editions Mir, Moscou, 1975.
SCHÖNBERG J. L.
[1] Metric spaces and completely monotone functions, Ann. Math. 39 (1939), 811-841.
SCHWARTZ L.
[1] Théorie des distributions, I, II, Paris, 1950, 1951.
SÉVASTIANOV B.
[1] Théorie des processus aléatoires branchus (en russe), Ouspekhi matem. naouk 6, nº 6 (1951), 47-99.
[2] Processus branchus (en russe), « Naouka », 1971.
SKOROKHOD A.
[1] Théorèmes limites pour processus aléatoires (en russe), Theoria veroïatnosteï i eïo primenenia 1 (1956), 289-319.
[2] Théorèmes limites pour processus aléatoires à accroissements indépendants (en russe), Theoria veroïatnosteï i eïo primenenia 2 (1957), 145-177.
[3] Théorèmes limites pour processus markoviens (en russe), Theoria veroïatnosteï i eïo primenenia 3 (1958), 217-264.
[4] Sur la dérivabilité des mesures associées aux processus aléatoires (en russe), Theoria veroïatnosteï i eïo primenenia 5 (1960), 45-53.
[5] Recherches sur la théorie des processus aléatoires (en russe), Izd-vo Kievsk. un-ta, 1961.
[6] Processus aléatoires à accroissements indépendants (en russe), Fizmathguiz, 1963.
SLOUTSKY E.
[1] Sur les fonctions aléatoires continues, intégrables et dérivables au sens stochastique. Comptes rendus Acad. Sci. 187 (1928), 370-372.
[2] Quelques propositions sur la théorie des fonctions aléatoires (en russe), Troudy Sr.-Az. un-ta, sér. matem. (5), 31 (1949), 3-15.
SPITZER F.
[1] A combinatorial lemma and its application to probability theory, Trans. Amer. Math. Soc. 82 (1956), 323-339.
[2] Principles of random walk, Princeton, 1964.
TCHENTSOV N.
[1] Convergence faible des processus aléatoires à trajectoires sans discontinuités de seconde espèce (en russe), Theoria veroïatnosteï i eïo primenenia 1 (1956), 154-161.
WATSON H. M., GALTON W.
[1] On the probability of the extinction of families, J. Antropol. Inst. 4 (1874), 138-144.
WIENER N.
[1] Differential space, J. Math. Phys. Mass. Techn. 2 (1923), 131-174.
[2] Extrapolation, interpolation and smoothing of stationary time series, N.Y., 1949.
YAGLOM A.
[1] Introduction à la théorie des fonctions aléatoires stationnaires (en russe), Ouspekhi matem. naouk 7, nº 5 (1955), 3-168.

INDEX DES MATIÈRES